AUG 16 1943

The Journal of Experimental Education

A periodical report of scientific investigations relating to child development, curriculum, learning, teaching, supervision, measurements, statistics, and experimental techniques.

Volume XI | JUNE, 1943 | Number 4

CURRICULUM CONSTRUCTION

CONTENTS

PAGE

Scholastic Prediction in a Teachers College: *Glenn W. Durflinger* ---------- 257

Effect of Size of Initial Food Serving on the Eating Efficiency of a Group of Preschool Children: *Marjorie Johnson Morse and Gertrude E. Chittenden*... 268

Value of Controlled Reading: Summary of Opinion and Research: *Arthur E. Traxler* ---------- 280

The Semantic Variations of Certain High Frequency Words in the Written Compositions of Eighth-Grade Pupils: *Harry K. Foster* ---------- 293

An Experiment With the Webster Complete Symbolization: *Ralph W. House*... 298

Constructing Equivalent Salary Schedules: *Joseph B. Hill* ---------- 302

$5.00 A YEAR | PUBLISHED QUARTERLY | $1.50 A COPY

Edited and Published by A. S. Barr, Professor of Education, University of Wisconsin, Madison, Wisconsin.

Entered as second-class matter October 17, 1938 at the post office at Madison, Wisconsin, under the act of March 3, 1879.

EDITORIAL BOARD

A. S. Barr, Chairman, Professor of Education, University of Wisconsin, Madison, Wis.

Carter V. Good, Professor of Education, University of Cincinnati, Cincinnati, Ohio. Editorially responsible for materials on learning, teaching and supervision, published each September.

J. Wayne Wrightstone, Assistant Director, Bureau of Reference, Research and Statistics, Board of Education of the City of New York, 110 Livingston Street, Brooklyn, New York. Editorially responsible for materials on curriculum construction, published each June.

Edward E. Cureton, Senior Personnel Technician, Personnel Research and Occupational Analysis Section, The Adjutant General's Office, 270 Madison Avenue, New York 16, N. Y. Editorially responsible for materials on measurements, statistics, and methods of experimental research, published each March.

Arthur T. Jersild, Professor of Education, Teachers College, Columbia University, New York City. Editorially responsible for materials on child welfare, guidance, and development, published each December.

CONTRIBUTING EDITORS

Gilbert L. Betts, Supervisor of Graduate Research in Education, Colorado State College, Fort Collins, Colorado.

William A. Brownell, Professor of Educational Psychology, Duke University, Durham, North Carolina.

Leo J. Brueckner, Professor of Education, University of Minnesota, Minneapolis, Minnesota.

Oscar K. Buros, Associate Professor of Education, Rutgers University, New Brunswick, New Jersey.

Otis W. Caldwell, General Secretary, The American Association for the Advancement of Science, Boyce Thompson Institute for Plant Research, Yonkers, New York.

Leslie L. Chisholm, Associate Professor of Education, State College of Washington, Pullman, Washington.

Herbert S. Conrad, Associate Professor of Psychology, College of Agriculture, and Research Associate, Institute of Child Welfare, University of California, Berkeley, California.

Stephen M. Corey, Professor of Educational Psychology, University of Chicago, Chicago, Illinois.

Robert A Davis, Professor of Education, Director of Bureau of Educational Research, University of Colorado, Boulder, Colorado.

Harl R. Douglass, Director of College of Education, University of Colorado, Boulder, Colorado.

Jack W. Dunlap, Associate Professor of Educational Psychology, University of Rochester, Rochester, New York.

Harold A. Edgerton, Assistant Professor of Psychology, Ohio State University, Columbus, Ohio.

Alvin C. Eurich, Professor of Education, Stanford University, Stanford University, California.

John C. Flanagan, Assistant Chief, Research Section, Medical Division, Office, Chief of the Air Corps, War Department, Washington, D. C.

Kai Jensen, Associate Professor of Education, University of Wisconsin, Madison, Wisconsin.

Harold E. Jones, Professor of Psychology, Director of Research, Institute of Child Welfare, University of California Berkeley, California.

Noel Keys, Professor of Education and Lecturer in Human Relations, University of California, Berkeley, California.

Edward A. Lincoln, Consulting Psychologist, Halifax, Massachusetts.

T. E. Newland, Pennsylvania State Department of Education, Chief of Special Education, Harrisburg, Pennsylvania.

C. W. Odell, Associate Professor of Education, University of Illinois, Urbana, Illinois.

Willard C. Olson, Professor of Education, Director of Research in Child Development, University of Michigan, Ann Arbor, Michigan.

W. E. Peik, Dean and Professor of Education, University of Minnesota, Minneapolis, Minnesota.

S. L. Pressey, Professor of Educational Psychology, Ohio State University, Columbus, Ohio.

Clarence E. Ragsdale, Associate Professor of Education, University of Wisconsin, Madison, Wisconsin.

H. H. Remmers, Director, Division of Educational Reference, Professor of Education and Psychology, Purdue University, Lafayette, Indiana.

Henry D. Rinsland, Professor of Education and Director of Educational Research, The University of Oklahoma, Norman, Oklahoma.

Robert T. Rock, Jr., Professor of Psychology, Head of Dept. of Psychology, Graduate School, Fordham University, New York City.

G. M. Ruch, Chief of Research and Statistical Service, U. S. Office of Education, Washington, D. C.

P. J. Rulon, Assistant Professor of Education, Graduate School of Education, Harvard University, Cambridge, Massachusetts.

Douglas E. Scates, Associate Professor of Education, Duke University, Durham, North Carolina.

David Segel, Educational Consultant, Specialist in Tests and Measurements, Federal Security Agency, U. S. Office of Education, Washington, D. C.

Paul W. Terry, Professor of Educational Psychology, University of Alabama, University, Alabama.

Helen Thompson, Clinic of Child Development, Research Associate, Yale University, New Haven, Connecticut.

Robert L. Thorndike, Associate Professor of Education, Teachers College, Columbia University, New York City.

Herbert A. Toops, Professor of Psychology, Ohio State University, Columbus, Ohio.

T. L. Torgerson, Professor of Education, University of Wisconsin, Madison, Wisconsin.

Helen M. Walker, Professor of Education, Teachers College, Columbia University, New York City.

Beth L. Wellman, Professor of Psychology, Child Welfare Research Station, State University of Iowa, Iowa City, Iowa.

Guy M. Wilson, Professor of Education, Boston University, Boston, Massachusetts.

Paul A. Witty, Professor of Education, Director of Psycho-Educational Clinic, School of Education, Northwestern University, Evanston, Illinois.

Ernest R. Wood, Professor of Education, New York University, New York City.

DEMOCRAT PRINTING COMPANY
MADISON, WISCONSIN

Journal of Experimental Education

Volume XI | June, 1943 | Number 4

SCHOLASTIC PREDICTION IN A TEACHERS COLLEGE

Glenn W. Durflinger
Santa Barbara State College
Santa Barbara, California

Introduction

Most colleges and universities of the nation have a published or unpublished study or two dealing with the prediction of college success. Most of these studies go no further than a simple correlational treatment with resulting interpretations. In the state of New York, where teachers college entrants are sharply restricted and each college and university admits its freshmen by quota, the colleges are obligated to set up the most accurate prediction formulas they can possibly develop. West of the Atlantic states entrance examinations are found much less frequently in the colleges.

New York state's selective admission is practiced in all of its teacher training institutions. Most of the larger universities over the nation have also been forced to select their students by some means. A study[1] of the teachers colleges in the East found that forty-six out of forty-seven institutions located along the Atlantic seaboard had in operation in 1934 some type of program for selective admission. Another more recent study covering a much wider area concluded with this statement,

> They (college catalogues) do seem to justify the conclusion that probably ninety percent of the institutions employ no means other than the traditional elimination of the more obvious misfits as they advance from entrance to the freshman class along the road toward graduation.[1]

Recent literature dealing with the improvement of teaching personnel contains many admonitions[2] that teachers colleges look to the future when a reduction in college students will be necessary and when only the best will be wanted as teachers.

Growing out of these very persuasive arguments that teachers colleges select their students more carefully and the few uncoordinated attempts at selection, an organization of teachers colleges developed which has as its primary purpose the construction, publication and distribution of tests and questionnaires. These materials can be and frequently are used as the basis for selecting the best teacher material. This organization claims, "The first national testing program among the higher institutions of learning was initiated in the fall of 1931 by the Teachers College Personnel Assocation."[3] Evidently this organization was the first in the field. Although it supplies tests for all higher institutions of learning, its initial purpose is to supply tests and norms from teachers college entrants to teachers colleges; the assumption being that the educational and intellectual components of the teachers college students are different from those of the liberal arts or university students. The number of colleges taking part in these annual programs by purchasing tests has varied from thirty-seven to eighty-three.

The Nebraska State Teachers College at Kearney is one of the colleges which participated in the 1939 testing program of the Teachers College Personnel Association. The

[1] Herbert D. Welte, "Present Status of Selective Admission in Teacher-Training Institutions in the Eastern-States Association of Professional Schools for Teachers, and Propaganda for Future," Problems in Teacher Training, (New York: Prentice-Hall, 1934) pp. 76–81.

[1] W. J. McConnell, "Application of the Principle of Selective Admission to Teachers Colleges and Its Significance for the Profession", The American Association of Teachers Colleges, *Nineteenth Yearbook* (Menasha, Wis.: George Banta Publishing Co., 1940) pp. 76–81.

[2] C. H. Judd, *Preparation of School Personnel*, (New York: McGraw-Hill Book Co., 1939) p. 93.

E. S. Evenden, "What is the Essential Nature of an Evolving Curriculum of a Teachers College," The American Association of Teachers Colleges, *Seventeenth Yearbook* (Menasha, Wisconsin: George Banta Publishing Co., 1938) p. 11.

Grayson N. Kefauver, "Needs and Challenges in Teacher Education," The American Association of Teachers Colleges, *Eighteenth Yearbook* (Menasha, Wisconsin: George Banta Publishing Co., 1939) pp. 80–81.

[3] J. D. Heilman, *The 1937 Report on the Cooperative Testing Program of the Teachers College Personnel Association* (Greeley, Colorado: Colorado State College of Education, 1937) p. 1.

tests used were those of college aptitude, elementary achievement, personal data and English. In addition to these tests the Bernreuter Personality Inventory and the Providence Music Test were administered to all entrants. These included 243 women and 140 men—a total freshman class of 383 students. For one reason or another 62 of this group, 30 women and 32 men, failed to take the entire battery of tests or did not complete the first semester's work and are not included in the group studied. The test results finally selected are those earned by 321 students, 213 women and 108 men.

Not all of the students in this group of 321 were anticipating teaching careers although considerably more than fifty percent report that teaching is the goal of their training. The college has set up special curriculums for pre-medicine, pre-nursing and pre-law students and likewise attempts to satisfy the higher educational needs of those desiring a liberal arts or general education.

The Data Collected

The tests were administered by members of the department of education faculty of the college. The students were informed of the fact that these test results were to be used in sectioning students in English and music classes. They were also placed on file in the registrar's office for the use of the faculty in individual guidance, personnel work or any use the members wish to make of them.

In addition to these data from the entrance examinations the first semester grades for each student were obtained from the registrar of the college. They were transmuted into grade point averages on the basis of the grade values as follows: A — 4 points, B — 3 points, C — 2 points, D — 1 point and F — 0 points.

The Nebraska state teachers college is not at present practicing selective admission. It is not using any formula developed elsewhere to investigate the efficiency of any method of college prediction. But the college administration and faculty are looking to the future when selective admission may not only be desirable but compulsory. It is far more sensible to investigate the possibilities and formulas for a particular college with the various tests it is using than to make no attempt at studying the possibilities of prediction. Other colleges the nation over are sounding the warning that selective admission is a desirable change, as well as one that will shortly be followed in all state supported teachers colleges.

The Zero Order Correlations. The relationships among the variables of the present study are indicated by means of coefficients of correlation. The zero order correlation coefficients were obtained through the use of the Pearson product moment formula. In using the correlation technique there are certain assumptions that are made which should be mentioned. These assumptions are briefly stated as follows:

1. The relationship between the two variables may best be described by a straight line equation.

2. This line is best determined by the method of least squares.

Table I

Zero Order Correlations and Their Probable Errors

r_{12}	M	.413 ± .054	r_{34}	M	.696 ± .033
	F	.509 ± .044		F	.764 ± .029
	T	.477 ± .029		T	.741 ± .017
r_{13}	M	.465 ± .051	r_{35}	M	.302 ± .059
	F	.512 ± .044		F	.475 ± .046
	T	.496 ± .028		T	.417 ± .031
r_{14}	M	.511 ± .048	r_{36}	M	.120 ± .064
	F	.506 ± .044		F	—.080 ± .056
	T	.508 ± .028		T	—.013 ± .038
r_{15}	M	.229 ± .062	r_{37}	M	.470 ± .051
	F	.355 ± .050		F	.475 ± .046
	T	.313 ± .034		T	.474 ± .029
r_{16}	M	.043 ± .065	r_{45}	M	.293 ± .059
	F	—.117 ± .056		F	.501 ± .045
	T	—.063 ± .037		T	.431 ± .031
r_{17}	M	.294 ± .059	r_{46}	M	.124 ± .064
	F	.378 ± .050		F	—.057 ± .056
	T	.350 ± .033		T	.004 ± .038
r_{23}	M	.679 ± .035	r_{47}	M	.460 ± .051
	F	.757 ± .030		F	.513 ± .044
	T	.731 ± .018		T	.496 ± .028
r_{24}	M	.841 ± .019	r_{56}	M	.104 ± .064
	F	.848 ± .013		F	—.157 ± .055
	T	.846 ± .011		T	—.070 ± .037
r_{25}	M	.221 ± .062	r_{57}	M	.247 ± .061
	F	.492 ± .045		F	.424 ± .048
	T	.401 ± .032		T	.365 ± .033
r_{26}	M	—.025 ± .065	r_{67}	M	—.006 ± .065
	F	—.069 ± .056		F	—.175 ± .055
	T	—.054 ± .038		T	—.119 ± .037
r_{27}	M	.394 ± .055	M men		
	F	.491 ± .045	F women		
	T	.458 ± .030	T total group		

1. Grade Point Average.
2. Intelligence.
3. Achievement.
4. English.
5. Music.
6. Personality.
7. Personal Data.

3. Standard scores are comparable.

To test the linearity of the relationships among the variables in this study a scatter-diagram was made for each two of the variables. The scatter-diagrams represented the relationships as linear rather than curvilinear. Lindquist says,

> If the curvilinearity is not so marked that it is not immediately apparent upon inspection of the scatter-diagram, the student need have no fear that the use of product-moment correlation technique will lead to any serious error.[1]

It will be observed from Table I that the highest zero order correlation is between the variables—intelligence and English. With the criterion measure—grade point averages—all correlations are significant except the one with the personality trait, neurotic tendency. A noticeable sex difference is found between the correlations of the criterion with intelligence, achievement, music and personal data.

One of the important contributions of this study is indicated in this summary table. The elementary subject matter tests correlate practically as highly with college grades as do high school subject matter tests. Douglass,[2] Segel,[3] and Wagner[4] found median correlations between tests of high school subject matter and college grades of about .55. The present study found correlations between tests of elementary school subject matter and college grades for the total group to be .50 for general achievement and .51 for English. The multiple correlation of these two achievement tests with college grades is $r_{1(34)} =$ .533 for men and .542 for women. These correlations give considerable weight to the assertion that an achievement test covering the elementary school subject-matter is, for all practical uses, almost as valuable in predicting college grades as is an achievement test covering high school subject matter. More studies, such as the present one, are needed to test the truth of this statement.

The question naturally arises as to the reason for the elementary school achievement tests being practically as valuable in prediction as are the secondary school achievement tests. The answer probably can be given under three points. (1) The fundamentals of education are obtained by practically all children in the elementary school. The secondary school curriculum merely increases the depth of learning on this foundation. It is the degree to which the child learns his fundamentals that determines to a considerable extent whether or not he will be a successful student either in high school or in college. (2) The habits of study which are so closely related to success in college are established primarily in the grades where and when the fundamentals are learned. (3) Success in college academic work as well as successful achievement in the elementary school subjects depends upon a common general factor such as Spearman's "g".

[1] E. L. Lindquist, *A First Course in Statistics*, (Chicago: Houghton–Mifflin Company, 1938) pp. 174–175.

[2] Harl R. Douglass, "The Relation of High School Preparation and Certain Other Factors to Academic Success at the University of Oregon," (Eugene, Oregon: University of Oregon Press, 1931) p. 50.

[3] David Segel, "Prediction of Success in College", *U. S. Office of Education, Bulletin No. 15* (1934) p. 61.

[4] Maxie Earle Wagner, *Prediction of College Performance*, (Buffalo, New York: The University of Buffalo, 1934) p. 64.

The Prediction Value of Combinations of Tests

The Multiple Correlations. The zero order correlations alone show nothing more than an existing relationship. Since the relationship with grades in college is the measure to be predicted, the technique of multiple correlation will add to an understanding of how important is each test's correlation with the criterion. Table II gives the multiple correlations.

The intelligence test is somewhat less indicative of college grades for the men than it is for the women; the zero order correlations being .413 and .509 respectively. By adding the achievement variable multiple r for men is raised to .484 and to .545 for women. Adding the English test raises the coefficients to .539 and .550. The English test as well as the achievement test raises the multiple r for men more than for the women. Further additions of the music, personality and personal data tests raise these multiple correlations an insignificant amount. The highest multiple r obtainable where only two tests from the battery are correlated with the criterion is found by using the English and the achievement tests. For men this correlation is .533 and for women .542. The final multiple r in which all six tests are used is .543 for men and .562 for women.

It appears significant that the multiple correlations do not run above .562 for either sex. The English test alone correlated with college marks gives a coefficient of .511 for

TABLE II

SUMMARY OF MULTIPLE CORRELATIONS

	r_{12}	$r_{1\,23}$	$r_{1\,24}$	$r_{1\,234}$	$r_{1\,2345}$	$r_{1\,23456}$	$r_{1\,234567}$
Men	.413	.484	.533	.539	.540	.542	.543
Women	.509	.545	.542	.550	.552	.556	.562
Total	.477	.524	.539	.540	.545	.548	.562

1. Grade Averages.
2. Intelligence.
3. Achievement.
4. English.
5. Music.
6. Personality.
7. Personal Data.

men and .506 for women. By adding the achievement test of the other elementary fields to the English test a multiple correlation is obtained which is almost as high as the multiple r which included all six tests. Even the inclusion of the intelligence test raises the multiple r only a small amount. It appears that a comprehensive achievement test of the elementary school field which includes reading and the elements of English composition and literature as well as arithmetic, science and the social sciences predicts college scholarship almost as well as a comprehensive high school achievement battery and better than an intelligence test alone.

The Significance of the Correlation Coefficients for Prediction. There are two appropriate methods of determining the significance of the zero order and multiple correlation coefficients in predicting college grades. The first to be used is the amount and frequency of displacement and the second is the relative size of the standard error of predicting true scores of grades. Table III indicates the degree of reliability that can be placed upon a prediction based on a correlation coefficient of .00. Only 26 percent of the predictions would be within one letter grade of accuracy and 2 percent would disagree four letter grades from those earned. On the other hand, entering the table with a multiple correlation coefficient of .55, which is conservative for the data presented in this study, out of a hundred possible grade averages predicted forty will be correct, forty three will be missed only one letter grade and only seventeen will be displaced two letter grades. There will be no displacements from the prediction as much or more than three letter grades in the five point grading system which obtained at the Kearney college.

Using the standard error of measurement it is possible to tell within what grade limits from prediction the grades will fall which

TABLE III

SUMMARY OF THE AMOUNT OF AGREEMENT AND DISAGREEMENT BETWEEN PREDICTIONS AND ACTUAL FRESHMAN MARKS[1]

r	Number of Divisions Displaced				
	0	1	2	3	4
.70	52	40	8	0	0
.60	46	40	14	0	0
.55	43	40	17	0	0
.50	40	40	20	0	0
.40	36	44	16	4	0
.00	26	40	24	8	2

[1] Charles W. Odell, *Statistical Method in Education* (New York: D. Appleton-Century Company, 1935) p. 181.

are one, two, or three standard deviations removed from the mean grade. The standard error of measurements is given by the equation $\sigma_{m1} = \sigma_1\sqrt{1 - r_{11}}$. Substituting in this formula the following known value for men $\sigma_1 = .80$ and for the women $\sigma_1 = .72$ with $r_{11} = .74$ for men and .71 for women the standard error values become as follows:

Men	Women
$\sigma_{m1} = .41$	$\sigma_{m1} = .38$

It is evident from these standard errors that approximately two thirds of the grades will fall within .4 of a grade step and approximately one third will fall beyond .4 of a grade step.

It is advantageous in this study to determine how near it is possible to predict, assuming that the grades were the true grades of the students. In other words within what limits is it possible to predict grades with the reliability and correlations already secured. It is possible to estimate very closely the correlations which should be obtained from completely reliable college grades and the tests. The estimate of such "true" grades can be made with less error than can the estimate of actual grades. In the following formula σ_1 is the standard error of the obtained marks, r_{11} is the reliability coefficient for col-

lege marks; found to be .74 for the men and .71 for the women. r_{1x} is the correlation coefficient for grades and the test variables. Using the formula $\sigma_{\infty x} = \sigma_1\sqrt{r_{1I} - r_{1x}^2}$ and the zero order correlation of each variable as well as the multiple correlations of these variables with the college grade index the following table of standard errors of true scores is calculated:

TABLE IV

STANDARD ERRORS OF PREDICTING TRUE SCORES OF GRADES BY MEANS OF THE VARIOUS TESTS

	Men	Women		Men	Women
$\sigma_{\infty 2}$ =	.60	.48	$\sigma_{\infty 23}$ =	.57	.48
$\sigma_{\infty 3}$ =	.58	.48	$\sigma_{\infty 234}$ =	.55	.46
$\sigma_{\infty 4}$ =	.55	.49	$\sigma_{\infty 2345}$ =	.54	.46
$\sigma_{\infty 5}$ =	.66	.55	$\sigma_{\infty 23456}$ =	.54	.46
$\sigma_{\infty 6}$ =	.69	.60	$\sigma_{\infty 234567}$ =	.54	.46
$\sigma_{\infty 7}$ =	.65	.54	$\sigma_{\infty 1234567}$ =	.58	.45

∞ True college grades.
2. Intelligence.
3. Achievement.
4. English.
5. Music.
6. Personality.
7. Personal Data.

Using the intelligence test in predicting college grades, the grades of men can be predicted so accurately that approximately two thirds of them will fall within .60 of a letter grade from the "true" grade. For the women approximately two thirds of the grades given will fall within .48 of a letter grade of the "true" grade. Since three standard deviations above and below the mean include practically 100 percent of the group, it is safe to say that no "true" grade for the men will be displaced more than 1.80 of a letter grade from the one predicted by the tests. For the women the "true" grade will be within 1.44 letter grades of the one predicted. This is a more exact statement of the grade prediction than the one which used the Odell tables. The highest possible prediction for the men gives a $\sigma_{\infty x}$ of .54 with the use of the intelligence, achievement and English tests. For the women the same tests give a $\sigma_{\infty x}$ of .46.

While it is not the purpose of the present study to defend the use of any or all of the college entrance examinations in use by the college studied, it should be pointed out that any procedure which makes possible the prediction of average true college marks within .45 to .53—depending upon the sex—of a letter grade in sixty-seven percent of the cases is not to be considered lightly.

The Regression Equation. A more refined technique of predicting grades by giving due weight to each of the test variables in comparison with all the others is the regression equation. The regression equations for both sexes and the total group in standard score form as found in the present study are given below.

The regression equation for the men shows that the English test is the most significant one in the equation. The achievement test comes next and beyond these two the other tests have either insignificant or slightly negative weightings.

In the regression equation for the women the largest weighting is found in the achievement test. About two thirds as much weighting as this test offers is found in the intelligence test and about half as much in the English test. The other weights are so small as to be practically insignificant.

Summary. The following generalizations seem warranted from a study of these equations and the multiple coefficients as determined by these tests in the institution furnishing the data.

1. For the men the test that is most predictive of college grades is the English test.

2. The next most valuable one is the elementary achievement test.

3. All the other tests have relatively small predictive weightings.

Regression Equations

Men	$\bar{z}_0 = -.146z_1$	$+.224z_2$	$+.457z_3$	$+.060z_4$	$-.050z_5$	$+.021z_6$
Women	$\bar{z}_0 = +.160z_1$	$+.221z_2$	$+.117z_3$	$+.062z_4$	$-.055z_5$	$+.099z_6$
Total	$\bar{z}_0 = +.057z_1$	$+.216z_2$	$+.230z_3$	$+.067z_4$	$-.052z_5$	$+.082z_6$

0. Criterion: first term's grades.
1. Intelligence.
2. Achievement.
3. English.
4. Music.
5. Personality.
6. Personal Data.

4. For the women the elementary achievement test carries the greatest weight in prediction.

5. The weighting of the achievement test is followed in order by those of the intelligence, the English, and the personal data tests.

6. The other two tests used in this battery proved to be of almost insignificant value in prediction.

7. For the total group of both sexes the English and the elementary tests give almost as good prediction as the entire battery. The employment of a third or fourth variable from this group will not significantly increase the predictive effectiveness of the equation.

8. It is important that the two sexes be treated separately with their own regression equations for the highest validity in prediction.

The Factors Found in Achievement in College

There have been a number of problems arise in this study which require a more refined technique to reach a solution. The fact of the negative weighting which intelligence has with college grades in the regression equation, as well as the near zero weightings of the factors of music, personality and personal data, suggests a need to inquire into the factors which are included in the pattern of college grades. The best of the known techniques for making this type of study in factor analysis, which is the one used.

Factor Analysis by the Centroid Method. There are several methods of multiple-factor analysis, of which the one most widely used is Thurstone's centroid method. Factor analysis is designed to discover the independent factors needed to account for the scores in a group of intercorrelated variables. It is used in the present study specifically to ascertain the factor characteristics of college academic success and to find how heavily weighted with these factors is each of the six college entrance tests.

There are a number of references available which will explain the step by step procedure involved in this analysis so it will not be reported here. It is sufficient to say that the method explained and followed by K. J. Holzinger and H. H. Harman[1] is used in this study. The results of each major step are presented and interpretations made when and where desirable.

[1] Karl J. Holzinger and Harry H. Harman, *Factor Analysis*, (Chicago: The University of Chicago Press, 1941) Chapter VIII.

It might be explained at the outset that the personality test was excluded from this analysis. All of the zero order correlations with personality were low and all were within the range of three probable errors. Any findings based upon such low correlations would be highly questionable if not definitely misleading.

The original and final factor loadings as found by means of the centroid method are given for the two sexes and the total class in Tables V, VI, and VII.

In the final factor loadings for the men there appears a factor that is common to the criterion and the five tests. This factor is almost twice as large in the first four variables as it is in the music and personal data tests. The second factor is practically nonexistent in scholarship average, does not appear in music but is prominent in intelligence, achievement, English and personal data. Beyond this second factor a third one is found in the elementary, music and personal data tests but since it is not present in an appreciable amount in grades it might be ignored for the purposes of the present study. The size of the communality demonstrates the fact that the criterion of college success is less than half measured by the test battery. In other words, the factors brought out by these tests as common to them and college scholarship are not as large as the unique or still unmeasured factor in scholarship. This may mean that in a teachers college, such as this one, where there is so much emphasis laid upon the preparation of teachers through the professional courses and the professionalized subject matter courses, the professors grade their students not only upon achievement past and present, or demonstrated and suspected intelligence, but also to a considerable extent they grade them upon potential teaching success as the professors subjectively measure it.

Only two factors are found in college scholarship from the final factor loadings for the women. The first and by far the largest is common to all six variables in almost the same amounts. The English test, as was the case with the men, contains this common factor in the greatest amount. Beyond the one general factor there appears a group

TABLE V

FACTOR LOADINGS: MEN

Original Factor Loadings

Variables	Factors 1	2	3	4	Communality h^2	Uniqueness $1-h$
1	.593	—.330	—.102	—.038	.4723	
2	.830	.254	—.289	.023	.8374	
3	.810	.070	.056	—.014	.6643	
4	.892	.070	—.214	.015	.8466	
5	.390	—.191	.264	.039	.2598	
6	.572	.126	.288	—.026	.4267	

Final Factor Loadings (After Rotation)

1	.678	.060	.099	.021	.4734	.5266
2	.581	.703	.000	.074	.8373	.1627
3	.578	.477	.319	.036	.6646	.3354
4	.713	.570	.091	.077	.8474	.1526
5	.336	.000	.378	.068	.2604	.7396
6	.300	.354	.460	.000	.4269	.5721

TABLE VI

FACTOR LOADINGS: WOMEN

Original Factor Loadings

Variables	Factors 1	2	3	4	Communality h^2	Uniqueness $1-h^2$
1	.620	—.149	—.049	—.051	.4116	
2	.882	.242	.009	.010	.8368	
3	.838	.158	—.021	—.022	.7281	
4	.890	.220	.013	.013	.8409	
5	.615	—.221	.026	.027	.4284	
6	.625	—.246	.023	.024	.4522	

Final Factor Loadings (After Rotation)

1	.636	.089	.000	.000	.4124	.5876
2	.726	.548	.066	.069	.8366	.1634
3	.721	.454	.035	.036	.7284	.2716
4	.742	.531	.071	.073	.8423	.1577
5	.645	.020	.076	.079	.4284	.5716
6	.664	.000	.075	.088	.4540	.5460

TABLE VII

FACTOR LOADINGS: TOTAL GROUP

Original Factor Loadings

Variables	Factors 1	2	3	4	Communality h^2	Uniqueness $1-h^2$
1	.611	—.154	—.053	—.054	.4027	
2	.865	.300	.011	.011	.8384	
3	.829	.114	—.022	—.022	.7012	
4	.890	.212	.010	.011	.8372	
5	.543	—.245	.027	.028	.3564	
6	.608	—.226	.026	.027	.4222	

Final Factor Loadings (After Rotation)

1	.624	.111	.000	.000	.4019	.5981
2	.658	.629	.068	.068	.8385	.1615
3	.707	.445	.039	.039	.7008	.2992
4	.717	.559	.072	.073	.8374	.1626
5	.587	.000	.078	.079	.3569	.6431
6	.638	.044	.081	.083	.4224	.5776

factor in the tests of intelligence, achievement and English. The other factor loadings are so small that they may be neglected.

The communalities indicate there is a unique factor or factors in college scholarship that are not measured by these tests. Differing as they do from unity, they indicate that there is a uniqueness factor which for both men and women is greater than the communalities in variable one—college grades. For the men the communality is represented by .4734 and the uniqueness by .5266. These are .4124 and .5876 for the women. Comparing the two sexes, it is evident that the battery of tests contains factors which overlap those in college scholarship to a slightly greater extent for men than for women.

In these uniquenesses there are two components, unreliability and specificity. From the data on the reliability of the criterion the reliability of college grades for the men is .74 and for women .71. The unreliability factor for the data for men is .26 and for the data for the women is .29 from the equation $u^2 + r_{11} = 1$.[1] This leaves for both the men and women a factor in college grades which is larger than any secured through the analysis of the present study except the general factor.

Men		Women
.53	uniqueness	.59
.26	unreliability	.29
.27	specificity	.30

The following tables are a complete summary of the factors, loadings, communalities and percent of their totals which are found in grades. In the table for men 47 percent of the components of grades are found in the four factors, twenty-six percent is accounted for by unreliability and twenty-seven percent

TABLE VIII

SUMMARY OF FACTORS, LOADINGS, COMMUNALITIES AND PERCENT OF TOTALS FOUND IN GRADES FOR MEN

Factors	Loadings	Communality	Percent
1	.678	.4597	47
2	.060	.0036	
3	.099	.0098	
(a)[1]	.500	.2600	26
(b)[2]	.495	.2669	27
Total		1.0000	100

[1] Holzinger and Harman, *op. cit.* p. 15.

TABLE IX

SUMMARY OF FACTORS, LOADINGS, COMMUNALITIES AND PERCENT OF TOTALS FOUND IN GRADES FOR WOMEN

Factors	Loadings	Communality	Percent
1	.636	.4045	41
2	.089	.0079	
(a)[1]	.500	.2900	29
(b)[2]	.581	.2976	30
Total		1.0000	100

TABLE X

SUMMARY OF FACTORS, LOADINGS, COMMUNALITIES AND PERCENT OF TOTALS FOUND IN GRADES FOR TOTAL GROUP

Factors	Loadings	Communality	Percent
1	.624	.3894	40
2	.111	.0123	
(a)[1]	.500	.2500	28
(b)[2]	.590	.3186	32
Total		1.0000	100

[1] (a) represents unreliability.
[2] (b) represents specificity.

is a specific factor the nature of which is unknown. A similar interpretation applies to Tables IX and X. It is beyond the bounds of the present study to conjecture what is the nature of this specificity. The author suggests that it is a multifactor consisting of interest, methods of work and perhaps the instructor's subjective opinion as to the type of teacher the student will make.

At this point it is advisable to compare the factors found for the two sexes. In the first place a general or common factor seems to be present in both sets of loadings. The second factor for the women coincides with the second for the men except for the fact that test six—personal data—contributes to this factor for the men but not for the women. The respective weightings of this factor in college scholarship are about equal and both low. The third and fourth factors for the women are very small yet common to all but the criterion of college success. The third factor for the men on the other hand is found in scholarship and all other variables but intelligence. These weightings, however, are low when compared with those of the first two factors. The last factor for the men is found in all variables except personal data; although in negligible amounts.

This analysis by the centroid method shows the six variables form factor patterns as illustrated in Tables XI, XII, and XIII.

TABLE XI

FACTOR PATTERN PLAN FOR SIX VARIABLES MEN

Variable	F_g	F_1	S
1	a		c_1
2	a	b	c_2
3	a	b	c_3
4	a	b	c_4
5	a		c_5
6	a	b	c_6

TABLE XII

FACTOR PATTERN PLAN FOR SIX VARIABLES WOMEN

Variable	F_g	F_1	S
1	a		c_1
2	a	b	c_2
3	a	b	c_3
4	a	b	c_4
5	a		c_5
6	a		c_6

TABLE XIII

FACTOR PATTERN PLAN FOR SIX VARIABLES TOTAL GROUP

Variable	F_g	F_1	S
1	a		c_1
2	a	b	c_2
3	a	b	c_3
4	a	b	c_4
5	a		c_5
6	a		c_6

The factor pattern plans as found in the above tables show a general factor F_g running through all the variables. In addition to this general factor there is one group factor F_1 which for the men is found in four tests and for the women and total group is found in only three. In addition to these factors the centroid method indicates that there are specific factors S in each test which are not identified in this study.

The factor analysis carries the study of factors in the tests somewhat beyond the original purpose in using it. The present study is concerned with the structure of college grades as the tests used reveal it. As soon as it is learned that the group factors in the tests are not also found in freshmen grades the analysis has served its purpose.

CONCLUSIONS

The following conclusions seem warranted from these solutions of the problems:

1. The two elementary survey tests, English, and achievement, used in the battery which is studied, have enough factors in common with college freshmen grades that they are good tests for scholastic prediction.

2. The remaining four tests of the battery of six contain practically no factors in common with college grades which are not included in the English and achievement tests. Therefore, whatever values they may have for individual guidance, they are not effective scholastic prediction devices.

3. The single test having the highest predictive value is the English test.

4. A comprehensive elementary school achievement test is practically as valid in the prediction of college grades as a comparable high school achievement test or high school grades.

5. The value of each test in prediction differs for the two sexes. Therefore, for most dependable prediction a separate regression equation should be formulated for each sex.

6. Using the two best tests in this battery, the English and the achievement tests, college grades can be predicted so accurately that no student will be displaced more than two letter grades and approximately forty-three percent will be forecasted within one letter grade.

7. Women's grades can be predicted slightly more accurately than can men's grades.

8. The factors in college grades, which are also components of the tests, count for less than half of the factorial content of grades.

9. There is a specificity in college grades which composes from one-fourth to one-third of their factorial content that is not measured by any of the six college entrance tests.

An Interpretative Comparison. A comparison of these findings with those of a recent study by Goodman[1] reveals that both studies find a general factor. Recently Thurstone and Thurstone[1] have reported finding a general factor in their advanced studies with the Thurstone Primary Abilities. All of this

[1] Charles H. Goodman, *Ability Patterns of Engineers and Success in Engineering School*, (Pennsylvania State College, 1941).

[1] L. L. Thurstone and T. G. Thurstone, "Factorial Studies of Intelligence", *Psychometric Monographs* No. II, (Chicago: University of Chicago Press, Chicago, 1941).

arouses again an interest in the "g" factor of intelligence which has been a contention of Spearman for many years. The study of Goodman's finds that tests of primary abilities secure a multiple correlation coefficient of .51 with college grades. The summaries of Douglass,[2] Segel[3] and Wagner[4] show that the correlations using all types of high school achievement tests and college grades are usually between .50 and .60. Approximately the same as this are the reported correlations between intelligence tests and college grades. The present study reports a correlation of .539 for both sexes between an elementary achievement test and college grades. Four distinct types of mental measurements correlate with college grades giving almost identical results. If a "g" factor, such as Spearman's were postulated, the argument could well be made that this common factor accounts for the similarity among the correlations.

In his study of engineering students, who by the way were almost exclusively men, Goodman also reports the finding of a verbal factor and a reasoning factor, which he identified from Thurstone's seven primary abilities. This verbal factor can tentatively be identified in the present study as the first group factor found in tests of intelligence, achievement and English.

Goodman used only tests of the seven primary abilities and found through the multiple correlation technique an r of .51 with first semester grade averages. Only the tests of number, verbal, space, induction and reasoning contributed to this multiple r. The other correlations are so small that no combinations he could use would raise this correlation. This multiple r is somewhat smaller than that obtained in the present study so it is reasonable to conclude that the specific factor or factors accounting for the grades of engineers is at least as great as the specific factor or factors accounting for the grades of students attending the Nebraska teachers college.

Probable Value of the Findings. The findings of the present study should be of value because:

1. They reveal to the large group of interested teachers colleges the widely varying predictive efficiency of the tests of the Teachers College Personnel Association.

2. One important finding of this study will become increasingly significant as high schools in preparing students for college are directed by the implications of many research studies to make no college preparatory course prescriptions. This will result in high school content examinations becoming less and less valuable as predictive devices. The elementary achievement examination may supplant the high school achievement examination in predicting college grades for the present study indicates that at present they differ very little in that function.

3. They reveal the widely varying degrees of predictive efficiency, as between the sexes, noted in each of the regression equations. This should be of some service to college administrators and personnel workers in warning them against using the same predictive agents for the whole student body, upon which grave decisions are made that are vital to the future welfare of the student.

4. They verify the assertion that a combination of variables is superior to a single one in the prediction of academic success in college.

5. They indicate where economies may be made in the use of selective admission procedures. It is a common procedure among teachers colleges to use a comprehensive elementary achievement test to guide students in building up their backgrounds in elementary subject matter. Apparently the same test can be used also to predict college academic success with as satisfactory results as would accrue from the use of a high school achievement test or high school marks and with better results than an intelligence test alone would give.

6. They may be used as a basis for a further study to determine where the most promising teacher training recruits may be found in Nebraska.

7. They reveal considerable of the factorial nature of scholastic grades in the college studied.

Recommended Studies. The present study does not furnish and was not intended to furnish evidence that the college supplying the data for the study should immediately launch

[2] Douglass, *op. cit.* p. 50.
[3] Segel, *op. cit.* p. 61.
[4] Wagner, *op. cit.* p. 64.

out upon a selective admission program. The study is the first extensive one made in this institution upon the character of its freshmen entrants and the structure of the grades given. It does however set the stage for subsequent studies which in this institution and elsewhere will supply a sounder basis for selective admission as well as the giving of grades.

The following studies are recommended:

1. A study of the reversals of expectancy as determined by the regression equations found in this study.

2. A study of college grades similar to this one using tests of the following traits:

a. Interest.
b. Drive.
c. Study habits.
d. Attitudes.

3. A study of departmental grades to determine which can be predicted with the greatest accuracy and which with the least accuracy.

4. A study of the possibilities of a pattern of abilities and personality traits that is unique for teachers.

5. A study of the parts of sub-tests of the several tests published by the Teachers College Personnel Association to determine which ones contain the largest factors in the prediction of college grades or teaching success of teachers?

6. A study of the possibilities in using the T. C. P. A. tests for differential prediction i.e. to predict students' success in various subjects.

7. A study planned to determine whether the factor found as common to the variables in the present study of teacher training candidates is the same as the common factor Goodman found in his study of engineers and what is the nature of this factor.

8. The extent to which other culture and occupational groups reflect the same common factor in comparable mental measurements.

EFFECT OF SIZE OF INITIAL FOOD SERVING ON THE EATING EFFICIENCY OF A GROUP OF PRESCHOOL CHILDREN

MARJORIE JOHNSON MORSE and GERTRUDE E. CHITTENDEN
Iowa State College

INTRODUCTION

Few authors who write in the field of child development fail to discuss child feeding and many of their suggestions are based on research which can be classified roughly into three categories: investigations of nutritional requirements of children; studies of techniques used in influencing children to eat; and studies concerning the influence of such factors as weather, day of week, and specific foods on eating efficiency. There are other suggestions regarding size of initial food servings and size and number of additional servings that are not based on scientific evidence, but are merely the expression of opinion.

Authorities agree that the size of initial food serving has some effect upon the eating efficiency of preschool children. However, the exact size of serving having the most favorable effect on eating efficiency as well as the degree of efficiency attained are details which writers in the field have either treated ambiguously or left unstated.

Blanton (4, p. 45) says that the food

> should be daintily served and the portions not be too large,

and Thom (17, p. 68) says further that

> dainty serving of food goes a long way in arousing appetite.

The quantitative connotation given to the word "dainty" is not made explicit. Barnes (3, p. 8) reiterates that

> The child should eat a small helping of each kind of food at a meal, given second helpings, if desired, when all other food has been eaten.

Blatz and Bott (5, p. 53) italicize the phrase *small portions* in their statement:

> Small portions attractively served will do more to stimulate the child's appetite than any amount of direct persuasion.

Aldrich (1, p. 58) views the problem from the standpoint of economy as well as appetite stimulation. He says,

> Plates should not be piled high with food. It is advisable to put so little on the dish that a second helping will be welcomed by the baby and requested by the older child. The sight of too much food decreases appetite just as the sight of a little stimulates it.

Paradoxically, he adds,

> If a second helping is requested, it should be given reluctantly. Possibly a remark that he should not eat so much or a grumble at the grocer's bill may help to get across to the child the idea that obtaining of food is difficult.

Stuart (16, p. 197) evades any direct statement of his opinion by saying,

> Ofttimes merely presenting too large a serving destroys appetite.

Anderson (2, p. 57) states,

> Giving the child a plate containing small amounts of the foods you wish him to eat and permitting him to receive a second helping if he wishes is preferable to overloading the plate. Many health workers advocate the gospel of the clean plate as a step toward good eating habits. By this they mean small servings so that the child can eat all the food and secure the satisfaction that comes from a clean plate.

Blatz, Millichamp, and Fletcher (6, p. 88) maintain a neutral position by the statement:

> In determining the amount of food to be served in each portion the effect upon behavior as well as the food value must be taken into account.

Haxton and Wilcox (8, p. 105) are dogmatic in their belief in small servings, but are vague as to exact amounts when they say,

> Place a small amount of food on his plate and put very little milk in his cup.

Justin and Snyder (9, p. 61) request that college students enrolled in a child development course of study observe

> Servings of a moderate size followed by second servings rather than large original servings.

Rand, Sweeny, and Vincent (16, p. 231) go somewhat further in their statement:

> Experience indicates that small portions, with an opportunity for second servings, inspire a much greater desire for food than large portions and have the further advantage of helping the child to form the habit of finishing what is placed on his plate since it usually can be eaten without undue effort.

A search of scientific literature in the area of child feeding revealed not a single investigation dealing with the influence of size of food serving upon eating efficiency. Consequently, the writer felt that the present problem was one which needed to be investigated.

A second influence in the choice of this particular problem was the investigator's experience in two nursery schools which differed widely in the size of initial food servings given to children at the noon meal. Believing that this discrepancy of practice resulted from a lack of scientific evidence in support of one method or the other she chose a problem which promised to throw some light upon this subject. As a verification of her belief that there is a widespread disagreement among nursery schools regarding size of food servings the investigator made a survey of nursery school noon meal procedures by means of a questionnaire which was sent to one hundred nursery schools in the United States. The results of this survey, based on the thirty-nine questionnaires which were answered completely enough to be used, are presented in Tables I and II as additional evidence that size of food servings is a point of disagreement.

Nineteen nursery schools considered their initial servings to be the minimum amount needed by the child, while fourteen considered them the optimum.

A disagreement was also found among nursery schools regarding practices in serving second helpings of food. Twenty-nine nursery schools believe in the "clean plate rule" to the extent that first servings of all foods must be eaten before second servings are offered, while ten nursery schools offer second servings before the child has eaten all of his first servings. However, twenty nursery schools do not require that the child eat all of his dinner before dessert is given, while fifteen do make that requirement. In addition to the data presented in the table, one nursery school reported that occasionally dessert is served when the child has not eaten all of his dinner and four nursery schools stated that this is rarely done in their mealtime procedure.

INVESTIGATION

Description of Subjects

The subjects of the investigation were the sixteen children who ate the noon meal at the Iowa State College Nursery School in Ames, Iowa during the spring quarter of 1942. Their ages ranged from three to four years at the beginning of the ten weeks quarter. The median age was three years five and one-half months. All but five of the children had been enrolled in the nursery school previous to the period of investigation and were familiar with the nursery school mealtime situation.

Method of Procedure

Arrangement of Dining Room

During the noon meal at the nursery school the children were seated at two small tables at each of which was a nursery school teacher. The seating arrangement remained the same throughout the investigation. Plates of food were on the tables when the children came into the room. The food was prepared by college students, but the measuring of food into servings was done by a graduate student who supervised the food preparation and serving.

This person, who also served second helpings of food in the dining room, sat at a small table near the children's tables. The investigator and an assistant observer were in the room. Each sat about two feet from the table which she observed. A person recording amounts of food served sat on a high stool placed in a position which afforded a view of each child's plate.

Food for additional servings was on a tray which the server carried to the tables whenever the teachers indicated that the children

TABLE I

SIZE OF INITIAL FOOD SERVINGS GIVEN TO NURSERY SCHOOL CHILDREN

Food	Age in yrs.	Range	Mean	Median
Meat or Substitute (Tbsp.)	2	.5 to 5	1.69	1.50
	3	1 to 5	2.19	2.00
	4	1 to 6	2.00	2.50
Egg (Part of whole)	2	.25 to 1	1.64	1.50
	3	.333 to 2	2.19	2.00
	4	.5 to 3	2.85	2.50
Cooked Vegetable (Tbsp.)	2	.5 to 3	1.55	1.50
	3	1 to 6	2.15	2.00
	4	1 to 8	2.88	2.50
Raw Vegetable* (No. & sizes of pieces of carrot)	2	1 piece 2" x ¼" to 1 med. carrot		
	3	1 piece 2½" x ¼" to 1 med. carrot		
	4	1 piece 2½" x ¼" x ¼" to 1 med. carrot		
Milk (Cups)	2	.5 to 3	.99	1.00
	3	.5 to 3	1.07	1.00
	4	.5 to 3	1.07	1.00
Bread (No. of slices ¼" thick)	2	.25 to 1	.56	.50
	3	.25 to 1	.66	.50
	4	.25 to 1	.74	.50
Soup (Tbsp.)	2	1 to 12	4.76	4.00
	3	2 to 12	6.50	6.50
	4	2 to 12	8.15	8.00
Fruit (Tbsp.)	2	1 to 8	3.08	3.00
	3	2 to 8	3.79	3.50
	4	2 to 12	4.84	4.00
Dessert (Tbsp.)	2	1 to 8	3.15	3.00
	3	2 to 8	3.76	3.00
	4	2 to 12	4.50	3.00

* Carrot is the only example of a raw vegetable given, since few nursery schools gave number and size of servings of other raw vegetables accurately enough to be used.

TABLE II

NUMBER OF NURSERY SCHOOLS ANSWERING YES AND NO TO QUESTIONS REGARDING SERVING OF FOOD

Question	Yes	No
Must first servings of all foods be eaten before second servings are offered?	29	10
Is a limit placed on number of servings given?	21	18
Is dessert ever given when the child does not eat all of his dinner?	20	15
Are second servings of dessert given when second servings of dinner were not given?	30	9
Do you consider the first servings the optimum amounts which you feel necessary for the child?	14	19
Do you consider the first servings the minimum amount which you feel necessary for the child?	19	14

were ready to be served. Dessert was also served by the server; milk was poured at the children's tables, a half cup at a time. Menus which were chosen by the dietitian were taken from Lowenberg's (11) menus for children and were those that in the past had proven to be well liked by the children at Iowa State College. Every attempt was made to keep the dining room as attractive and pleasant as possible.

Description of Servings

The two methods of serving food were called Method I and Method II. In Method I the size of initial food serving was that offered ordinarily at the Iowa State College Nursery School. Since, for Method II a size of food serving was needed which would be sufficiently small that a child would perceive the difference in size between it and the larger serving, it was decided that the servings in

PLATE 1.

PLATE 2.

Method II should be one-third the size offered in Method I. Plates 1 and 2 illustrate the same foods served under each method.

Food was served in units and all measurements were level. The following amounts of food were considered one unit:

Meat	1 Tablespoonful
Bacon	⅓ Slice
Egg	⅓ Whole egg
Soup	2 Tablespoonfuls
Cooked Vegetable	1 Tablespoonful
Raw Vegetable	
Carrot	Piece ⅔″ x ½″
Celery	Piece ⅔″ x ½″
Cabbage	Piece ⅔″ x 1″
Lettuce	Piece ⅔″ x 1″
Sandwich or Toast	⅙ of a slice of bread ¼″ thick
Milk	⅓ Cupful
Dessert	2 Tablespoonfuls
Cookies	1 Cookie
Gingerbread	1 Small cake

Method I consisted of serving initially three units of meat or egg, cooked and raw vegetables, and sandwich or toast according to the day's menu. Soup and dessert were served in two and one-half unit portions, the nearest equivalent of one-third cupful, to facilitate serving. A dipper holding exactly three tablespoonfuls of food was used in serving all of the food of that sized portion. Standard one-third cup measures were used for measuring soup and dessert.

The following portions were considered three units of food, with the exception of soup and dessert:

Meat	3 Tablespoonfuls
Bacon	1 Slice
Egg	1 Whole egg
Soup	⅓ Cupful (considered 2½ units)
Cooked Vegetable	3 Tablespoonfuls
Raw Vegetable	
Carrot	Piece 2″ x ½″
Celery	Piece 2″ x ½″
Cabbage	Piece 2″ x 1″
Lettuce	Piece 2″ x 1″
Sandwich or Toast	½ of a slice of bread ¼″ thick
Milk	1 Cupful
Dessert	⅓ Cupful (considered 2½ units)

Method II consisted of serving one unit of meat, egg, cooked and raw vegetables and sandwich or toast according to the day's menu. The dipper used to serve soup held three tablespoons of liquid, so soup was served initially in one and one-half unit portions. Two and one-half units of dessert were served for it was found during the period preliminary to the investigation that the children showed a tendency to consider the meal complete following their one small serving of dessert. Since dessert often contained food elements not duplicated elsewhere in the day's menus, it was felt that a larger serving should always be given. Second servings of cookies or gingerbread were never offered in either method of serving, and only one unit was served initially with both methods. A standard tablespoon measuring spoon was used to measure that sized portion of food. A dipper holding exactly three tablespoons liquid measure was always used to serve soup. A standard one-third cup measure was used to serve dessert.

When Method I was in progress, additional servings of food in one unit portions were offered to the child when only one initially served food remained on his plate. Having eaten three tablespoonfuls of all but one food, he was given a choice of additional servings of food. Following his second serving, additional servings were offered when he had eaten all of any one food on his plate. Similarly, for Method II, additional servings were given when the child had consumed all but one of his food servings. Servings of foods other than the one remaining on his plate were served in one unit portions. When he had eaten one of these second food servings, an additional portion was placed on his plate. A choice as to whether or not he desired more food was given only after he had eaten three units of each food.

Eating Time

Thirty-five minutes were allotted for eating the main course of the meal. At the end of that time the adult who served second helpings removed the child's plate and served his dessert. When a ten minute interval had elapsed the dessert was removed and the meal was over. If the child had not drunk his cup of milk, he was allowed to remain at the table beyond the time limit to finish drinking it as this was felt to be a nutritional requirement which should not be waived. All foods remaining on the child's plate at the end of the time allowed were measured and subtracted from the total amount served.

Teacher-Child Contacts

Teacher-child contacts were kept as uniform as possible during the noon meal. The teachers who sat at the tables with the children were limited in their choice of guidance techniques. They were allowed to:

I. Make comments about food such as:
   A. "The carrots are crisp."
   B. "Isn't our food pretty today?"

II. Make indirect suggestions concerning amounts of food to be eaten. These include:
   A. Getting food into position to eat.
   B. Placing food in the child's fork or spoon.
   C. Bringing the child back into position to eat.
   C. Making comments such as:
      1. "It's eating time now."
      2. "Take turns with your egg and sandwich."
      3. "A drink of milk will help."
      4. "I wonder what our dessert will be."
      5. "You're ready for another bite now."

III. Make comments about the technique of eating. Examples are:
   A. "Tip your bowl."
   B. "Hold your spoon like this."
   C. "It would be easier to use your spoon."
   D. "Chew it first, then swallow."

The teachers were not allowed to:

I. Make direct suggestions concerning amount of food to be eaten. These include:
   A. Feeding the child.
   G. Making comments such as:
      1. "Finish your food."
      2. "Let's get this plate cleaned up now."

A week's practice period allowed the teachers who sat at the tables with the children to become familiar with the use of these suggestions before the investigation started.

Period of Investigation

The investigation covered a ten weeks period. The children were divided into two groups, each of which ate lunch at the nursery school for a period of five weeks. A practice period of two weeks preceded the actual investigation. During this time both methods of serving were used, and the persons cooperating in the mealtime procedure became thoroughly acquainted with the methods.

Since some of the children had never eaten at the nursery school, a week preceding the investigation was allotted for making an adjustment to the nursery school mealtime situation. Method I was used during the first group's initial adjustment period, and to offset any effect this preliminary adjustment might have on eating efficiency under the method immediately following it, Method II was used during the second group's initial adjustment period. Following this period the methods were alternated at one week intervals.

Weekly menus were repeated with each method of serving. Thus, each child was served the same menu twice during the period of investigation. Menus used appear in the Appendix.

An outline of the period of investigation appears in Table III.

Index of Efficiency

The index of efficiency was used originally by McCay (13) and Macleod (12) at New York State College of Home Economics, Cornell University. Two components of efficiency, eating time and amount of food eaten, were used. The formula is:

Efficiency score = 100 — units of time + units of food.

TABLE III

WEEKLY DISTRIBUTION OF MENUS AND METHODS

	First Group					Second Group				
Week	1.*	2.	3.	4.	5.	6.*	7.	8.	9.	10.
Menu	A	B	C	C	B	A	B	C	C	B
Method	I	I	II	I	II	II	II	I	II	I

* Initial Adjustment Period.

The units of time were one minute in length and the food units were as previously described.

Method of Observation

The investigator and another person who had established reliability with her made the daily observations during the noon hour. Each of them observed the children seated at one table. During the period of investigation the investigator recorded observations for ten children and the assistant, for six children. Observations were noted on a large chart attached to a clipboard which the observers held in their laps. A stop watch and a pocket watch were held in the left hands of the observers as they recorded.

The following items were recorded:

1. *Starting time.* The pocket watches were set at twelve o'clock when the first child was seated at the table. The stop watches were also started and the time each child sat down at the table was recorded to the nearest fifteen second interval. This was considered starting time.

2. *Stopping time.* The time the child stood up from his chair preparatory to leaving the dining room after having eaten his meal was considered stopping time. This was checked to the nearest minute. If he had not completed his meal, stopping time was checked when the time allotted for eating his meal had elapsed.

3. *Time between seating and attack.* The moment the child put the first bite of food into his mouth was considered time of attack. Since the stop watches were put into action when he was seated at the table, it was possible to check attack to the nearest fifteen second period. The time interval between starting time and the moment the child placed the first bite of food in his mouth was called "time between seating and attack." This was observed because it was felt that the time between seating at the table and attack on food might be a measure of his interest in eating.

4. *Time out.* The time consumed in some activities was considered irrelevant to eating time so the time spent in such activities, to the nearest fifteen seconds, was marked on the chart with a straight line and subtracted from the total eating time. Time out was denoted for two activities:

a. Obtaining second servings of food. The stop watch was started the moment the tray bearing second servings touched the dining table at which the child to be served sat and was stopped when the tray was lifted.

b. Obtaining napkins, silverware, or any act requiring that the child leave the table other than removal from the group for undesirable behavior. This time was checked from the moment the child got up until he sat down again.

Removal from the table for undesirable behavior was not considered time out since it was felt that such behavior should affect the efficiency score. When the child was removed from the dining room he was seated at a small table just outside the door of the dining room where he could still be observed by the investigator.

5. *Total eating time.* This was considered as the period beginning with the time the child placed the first bite of food in his mouth and ending at the time the child stood up to leave the table at the end of the meal, with time out intervals subtracted.

Observer Reliability

Reliability was established between the investigator and a student trained in child development. This was done in a two weeks period previous to the investigation, during which time both observers recorded data for the same children. The four items checked on the chart were considered individually when establishing reliability.

The two observers agreed on starting time within thirty seconds 97.77 per cent of the time. Perfect agreement was obtained on time of attack on food and stopping time. A one minute interval was allowed for agreement, since the stopping time was checked to the nearest minute. The agreement on total eating time was 90.00 per cent within a range of one minute. The total time out agreement was 97.75 per cent which was an agreement on the number of fifteen second intervals checked at the same time by both observers. The two observers agreed 86.75 per cent of the time on length of observation and each checked the same intervals 93.26 per cent of the time.

TABLE IV

AGREEMENT OF TWO OBSERVERS

Observations	Agreements	Disagreements	Percentage
Starting Time	44	1	97.77
Stopping Time	50	0	100.00
Total Time Out	240	5	97.95
Length of Time Out Period	72	11	86.75
No. of Time Out Periods	83	6	93.26
Attack on Food	36	0	100.00
Total Eating Time	45	5	90.00

CHILDREN'S REACTIONS TOWARD THE INVESTIGATION

Few comments were made by the children either in regard to the observers or difference in size of food serving.

One child asked,

What are you doing with the watches?

He was given the answer:

To see how long it takes to eat lunch.

This apparently was a satisfactory answer, for the question was not repeated. Another child came into the dining room, glanced at the investigator holding the watches and asked—

What time is it?

The first day the very small sandwiches (one-third the usual size made at Iowa State College Nursery School) were served, a child picked up his sandwich, then looked at the teacher seated at his table and smiled. Later, after he had been given several additional servings following his small initial serving, he remarked,

I've had a lot of seconds today.

When Method II was in effect one small boy looked at the server and said,

I'd like a whole big sandwich.

Bacon was served one day when Method I was used. A child commented,

Look at my long piece of bacon.

That same day when a child was asked if she cared for another piece of bacon she answered,

Yeah, a long piece.

When the menu was repeated using Method II, a child picked up her piece of bacon, looked at it, and announced,

At home I get a long piece of bacon like my Mommy and Daddy.

Still another child summed up his attitude toward Method II in the trite phrase,

We don't have much to eat today, do we?

This was the total number of children's comments concerning the investigation which were recorded by the investigator.

RESULTS

EFFICIENCY SCORES UNDER THE TWO METHODS

In determining the effect of the two sizes of initial food servings on eating efficiency, data for thirteen children were used. Only those scores were used which were matched in menu and method of food serving, that is, a child's score was disregarded in comparing the results of the two methods unless he had eaten the same menu with both methods of serving. Unless a child had at least four matched scores, none of his scores were used in determining the effect of size of initial food serving on eating efficiency. Thus, the scores of two children who had been absent during a part of the investigation, and of one child who was dropped from nursery school before the period was completed were excluded.

A mean efficiency score was obtained for each child for Methods I and II.

Nine of the thirteen children had higher mean efficiency scores when they were given small initial food servings, while four ate more efficiently when large servings were given initially.

TABLE V

MEAN INDIVIDUAL EFFICIENCY SCORES FOR METHODS I AND II

Subjects	Mean Efficiency Scores Method I	Method II
K. A.	88.81	87.63
B. B.	91.83	84.21
B. L. C.	80.60	83.05
D. D.	83.95	86.45
D. D.	87.50	87.79
J. F.	82.83	90.75
V. H.	83.81	85.00
M. L. H.	76.06	78.50
P. J.	93.50	83.06
J. J.	89.78	86.72
C. L.	76.13	80.63
J. A. S.	86.45	88.50
N. T.	82.80	90.60

TABLE VI

COMPARISON OF MEAN EFFICIENCY SCORES FOR METHODS I AND II

Method	Range of Scores	Mean Score	Standard Deviation	Difference Between Mean Scores
I	76.06 to 93.50	84.93	5.2	0.70
II	78.50 to 90.75	85.61	2.9	

One child, N. T., raised her eating efficiency score by 8.60 points when small food servings were given initially.

Conversely, another child, P. J., lowered his score by 10.44 points when given small initial food servings.

Both the lowest and highest mean efficiency scores were obtained when large food servings were given initially.

Although a higher mean efficiency score was found when the children were served small portions of food initially, the difference between the mean scores of the two methods was not statistically significant when Fisher's t[1] test was used to determine the significance of the difference between the means. The value of t was .4607. An absolute value of t equal to .4607 would be found in between sixty and seventy per cent of random samples of this size. Only differences which would occur in ninety-five per cent of the cases, that is, where the value of t was significant at the five per cent level of confidence, were considered statistically significant by the investigator.

AMOUNTS OF FOOD EATEN UNDER THE TWO METHODS

The efficiency score was broken down into its components, eating time and amount of food eaten, in order to determine whether the method of serving influenced them.

[1] The significance of the difference between the means was determined by using the t test suggested in Lindquist, E. F. Statistical Analysis in Educational Research. p. 54. New York. Houghton Mifflin Co. 1940.

$$t = \frac{M_a - M_b}{\sqrt{\frac{\Sigma d^2}{n(n-1)}}}$$

TABLE VII

MEAN NUMBER OF FOOD UNITS EATEN FOR METHODS I AND II

Subjects	Food Units Eaten Method I	Method II
K. A.	29.88	28.00
B. B.	24.08	21.25
B. L. C.	20.50	21.30
D. D.	24.50	20.10
D. D.	21.25	22.66
J. F.	22.75	28.75
V. H.	21.50	23.75
M. L. H.	19.44	21.53
P. J.	21.13	30.75
J. J.	29.31	27.25
C. L.	18.25	21.38
J. J.	23.80	27.30
N. T.	24.10	23.70

Eight of the thirteen children ate a larger number of food units when small servings were given initially, while five ate a larger number of food units when large initial servings were given.

Again Method I presents the greater range of number of food units eaten. The children ate the largest and the smallest amount of food recorded when they were given large initial food servings.

The children ate more food when it was served in small portions. The difference between means of the two methods, which is 1.33 is not a statistically significant difference as the value of t is 1.25. An absolute value of t equal to 1.25 would be found in about twenty per cent of random samples of this size, that is, a difference of 1.33 food units or more would be found in only eighty per cent of random cases of this size.

TABLE VIII

COMPARISON OF NUMBER OF FOOD UNITS EATEN FOR METHODS I AND II

Method	Range of No. of Food Units Eaten	Mean Score	Standard Deviation	Difference Between Mean Scores
I	18.25 to 29.88	23.11	3.3	1.33
II	20.10 to 28.75	24.44	3.2	

Total Eating Time Under the Two Methods

Nine of the thirteen children ate in a shorter time when they were given small initial food servings. One child, N. T., lessened her mean total eating time by 8.20 minutes. Four children required a longer time to eat their noon meal when given small initial food servings. The largest difference was required by D. D. who stayed at the table 4.79 minutes longer when given small portions of food.

Table IX

Mean Total Eating Time in Minutes for Individuals for Methods I and II

Subjects	Total Eating Time in Minutes	
	Method I	Method II
K. A.	42.31	40.38
B. B.	32.25	37.04
B. L. C.	39.90	35.65
D. D.	40.55	33.65
D. D.	33.75	34.71
J. F.	40.13	38.00
V. H.	37.63	35.50
M. L. H.	43.38	43.06
P. J.	27.83	27.69
J. J.	39.41	40.53
C. L.	42.25	40.75
J. A. S.	37.45	38.80
N. T.	41.30	33.10

The lowest figure in the range of total eating time is found when the children were served by Method II, and the highest was found when Method I was in effect.

The difference between the total eating time means is 1.48 minutes and, although less time was required for eating when small food servings were given, the difference is not statistically significant, for the value of t is 1.54. An absolute value of t equal to 1.54 would be found in between ten and twenty per cent of random samples of this size.

Table X

Comparison of Total Eating Time in Minutes for Methods I and II

Method	Range of Total Eating Time	Mean Eating Time	Standard Deviation	Difference Between Mean Eating Times
I	27.83 to 43.38	38.32	4.4	1.48
II	27.69 to 43.06	36.84	3.3	

Time Between Seating and Attack on Food for the Two Methods

Seven children waited longer to start eating their food after being seated at the table when Method I was in effect, while for six children the delay was longer for Method II.

Table XI

Mean Time Between Seating and Attack in Minutes for Individuals for Methods I and II

Subjects	Time Between Seating and Attack in Minutes	
	Method I	Method II
K. A.	0.28	0.41
B. B.	0.29	0.63
B. L. C.	0.54	0.40
D. D.	0.30	0.40
D. D.	0.29	0.17
J. F.	0.33	0.15
V. H.	0.56	0.19
M. L. H.	0.25	0.34
P. J.	0.75	0.25
J. J.	0.91	1.78
C. L.	0.31	0.25
J. A. S.	0.35	0.70
N. T.	0.75	0.50

The greatest range in time was found when the children were given small initial servings of food.

The children waited longer after being seated to attack their food when Method II was in effect. This is the only measure of efficiency which did not show an advantage for Method II.

The difference between the mean time in minutes between seating and attack on food for both methods was 0.02. This difference was not statistically significant, as the value of t is 0.19. This value would be found in between eighty and ninety per cent of random samples of this size.

Time between seating and attack on food seemed a poor measure of the relative merits of the two methods. Some of the children

TABLE XII

COMPARISON OF MEAN TIME BETWEEN SEATING AND ATTACK IN MINUTES FOR METHODS I AND II

Method	Range of Time	Mean Time	Standard Deviation	Difference Between Mean Times
I	0.25 to 0.91	0.45	.2	0.02
II	0.15 to 1.78	0.47	.5	

obviously waited for certain preliminary activities before beginning to eat regardless of size of food serving. The teachers helped push the children's chairs into position and tucked their napkins in at the neck openings of dresses and suits. These acts obviously delayed the children's attack on food.

AMOUNT OF FOOD EATEN, LENGTH OF EATING TIME, AND EFFICIENCY SCORES FOR EACH DAY OF THE WEEK

A difference was found between the length of eating time, amount of food eaten, and efficiency scores for the different days of the week. In determining the daily mean amounts of food eaten, eating time in minutes, and efficiency scores all the scores of each subject in the investigation were used.

This was also evident in this investigation, for eating was accomplished in the shortest length of time on Wednesdays, and Wednesday's mean efficiency score was the second largest of the daily mean scores.

The fact that a higher number of food units was eaten on Tuesdays could be accounted for by the fact that the Tuesday meal for both Menus B and C contained bacon, toast, and a cream soup. Not only were these well liked foods, but they were easily eaten, for the bacon and toast were eaten with the fingers, and the soup was drunk from a cup. One child who had a motor difficulty and who often remarked, "I'm tired," before the meal was over, usually ate a larger number of food units on Tuesdays.

TABLE XIII

DAILY MEAN NUMBER OF FOOD UNITS EATEN, LENGTH OF EATING TIME IN MINUTES, AND EFFICIENCY SCORES

	Monday	Tuesday	Wednesday	Thursday	Friday
Mean No. of Food Units Eaten	22.84	26.56	23.15	23.78	24.78
Mean Eating Time in Min.	37.91	36.41	35.93	38.95	39.10
Mean Efficiency Score	83.41	89.68	86.35	85.01	85.64

The largest number of food units was eaten on Tuesday, and the smallest on Monday. Eating was accomplished in the shortest time on Wednesday, and the longest on Friday. Highest and lowest efficiency scores were made on the same days, Tuesday and Monday respectively, that the largest and smallest number of food units were eaten.

These findings compare favorably with those of McCay (13) who stated that lowered eating efficiency on Tuesday was a result of Tuesday's being the first day of school after the week-end. Comparably, on Monday, the first day of school at the Iowa State College Nursery School following the week-end, there were a longer eating time, a fewer number of food units eaten and the lowest mean efficiency score of the week. She also reported that Wednesdays brought higher efficiency scores.

TABLE XIV

MEAN EATING TIME IN MINUTES, MEAN NUMBER OF FOOD UNITS EATEN, AND MEAN EFFICIENCY SCORES FOR ALL SUBJECTS

	Mean	Standard Deviation
Eating Time	37.53	5.6
Foods Units Eaten	24.26	5.8
Efficiency Score	86.73	7.5

The figures in the foregoing table were computed from all scores of all children.

An example of the amount of food in a sample menu which would total twenty-four and one-fourth units (the nearest to the mean number of food units which could be measured accurately) is given in the following table:

TABLE XV

AMOUNT OF FOOD REPRESENTING THE MEAN NUMBER OF FOOD UNITS EATEN

Food	Amount	Food Units
Creamy Egg	1⅓	4.00
Scalloped Tomato	⅓ C.	2.50
Lima Beans	4¼ T.	4.25
Lettuce	1 piece 3" x 2⅝"	3.00
Peaches	⅓ C.	2.50
Sandwich	1⅓	4.00
Oatmeal Cookie	1	1.00
Milk	1 C.	3.00
Total		24.25

RELATION BETWEEN AMOUNT OF FOOD EATEN AND LENGTH OF EATING TIME

Amount of eating time in minutes was correlated with number of food units eaten. The coefficient of correlation was 0.14. After a correlation diagram was constructed, it was noted that there was an apparent tendency for the means of the columns to fall along a curved rather than a straight line.

A scatter diagram of eating time and food units was made. Linear regression was tested, and Z was found to equal .201 which is not significant. A significant departure from linearity would occur on the five per cent level when Z equals .270.

Thus, the coefficient of correlation is an adequate measure of the relationship between length of eating time and amount of food eaten, although the coefficient, .137, is not significant at the five per cent level of confidence.

This finding indicates that there is some tendency for children who stay at the table a long time to eat more food, but from the scatter diagram it was evident that some children who stay at the table a long time eat less food.

It is impossible for the writer to either agree or disagree, on the basis of her findings, with the findings reported by Dunshee (7) and Macleod (12) whose reports were contradictory, the former stating that children who sit at the table longer tend to eat less food, and the latter finding a significant relationship between food intake and eating time.

APPLICABILITY OF THE RESULTS OF THE INVESTIGATION

Although it was found that children eat more food in less time when given one tablespoon serving initially, the amount of time and energy consumed in serving additional servings may not be justifiable in most nursery schools since the difference between the two methods of serving was not statistically significant.

In light of the findings of the investigation, the following suggestions are made for the mealtime procedures at Iowa State College Nursery School:

1. Serve two units of all foods initially. These servings would be small enough that success in eating all the food could be achieved easily, and they would be large enough that the amount of time spent in giving additional servings would be reduced to the extent that both servings and recording of amounts of additional servings could be done by one person.

2. Allow a choice of additional servings of all foods when the child has eaten one-half of his initial food servings.

3. Limit the amount of time allowed for eating the dinner to forty-five minutes, with the last ten minutes of that time allotted for dessert.

4. Use as few direct guidance techniques in the eating situation as possible, for when their use was partially controlled it was found that the children ate as much food in a comparable length of time as they did before the start of the investigation.

SUMMARY

The investigation was conducted at the Iowa State College Nursery School at Ames, Iowa for the purpose of determining the effect of size of initial food servings on the eating efficiency scores obtained during the nursery school mealtime. Sixteen subjects furnished data for the investigation.

Two mealtime procedures were used alternately for four weeks. One method consisted of serving initially three tablespoonfuls or comparable sized portions of all foods on the menu. The other method consisted of serving initially one tablespoonful or comparable sized portions of all foods on the menu. As many additional servings as the child desired were given in both methods. The weekly menus were repeated with each method of serving. Teacher-child contacts during the noon meal were made as uniform as possible.

Two components of efficiency, the time taken for the child to eat the meal and the

amount of food eaten, were measured in units and an efficiency score obtained for each child for each meal. The length of the time interval between the moment the child sat down at the table and the moment he put the first bite of food in his mouth was also recorded.

The following results were obtained:

1. Higher efficiency scores were obtained when one tablespoon of each food was served initially. The difference between the mean scores of the two methods was not statistically significant.

2. The children ate more food and in a shorter length of time when they were served a small amount of food initially, although the difference between the two methods was not statistically significant.

3. The time between seating and attack on food was slightly longer when small amounts of food were served initially, although the difference was not statistically significant.

4. The largest number of food units was eaten and highest efficiency score obtained on Tuesday: the fewest number of food units was eaten and lowest efficiency scores obtained on Monday. On Wednesday the mean eating time was shortest, and on Friday it was longest.

5. A very slight positive, though statistically insignificant, correlation of .137 was found between number of food units eaten and length of eating time.

6. The mean eating time of all the subjects was 37.53 (S.D. 5.61), the mean number of food units eaten was 24.26 (S.D. 5.82), and the mean efficiency score was 86.73 (S.D. 7.51).

Literature Cited

1. Aldrich, Charles Anderson. Cultivating the Child's Appetite. p. 58. New York. The Macmillan Co. 1932.
2. Anderson, John E. Happy Childhood. p. 57–58. New York. D. Appleton–Century Co., Inc. 1933.
3. Barnes, Mary Frances Hartley. Feeding the Child from Two to Six. p. 8. New York. The Macmillan Co. 1928.
4. Blanton, Smiley and Blanton, Margaret Gray. Child Guidance. p. 45. New York. The Century Co. 1927.
5. Blatz, William E. and Bott, Helen. Parents and the Preschool Child. p. 53. New York. William Morrow. 1929.
6. Blatz, William E., Millichamp, Dorothy and Fletcher, Margaret. Nursery Education. p. 88. New York. William Morrow and Co. 1935.
7. Dunshee, Marion Elizabeth. A Study of Factors Affecting the Amount and Kind of Food Eaten by Nursery School Children. Child Development. 2: 163–183. 1931.
8. Haxton, Jennie N. and Wilcox, Edith. Step by Step in the Nursery School. p. 105. Garden City, New York. Doubleday, Doran and Co., Inc. 1936.
9. Justin, Florence and Snyder, M. Eunice. Directed Observation in Child Development. p. 61. New York. Houghton Mifflin Co. 1938.
10. Lindquist, E. F. Statistical Analysis in Educational Research. p. 54. New York. Houghton Mifflin Co. 1940.
11. Lowenberg, Miriam E. Your Child's Food. 2nd ed. p. 149–203. New York. Whittlesey House. 1934.
12. Macleod, Ila. Responses of twelve nursery school children to the noon meal at nursery school during the year. Unpublished M.S. Thesis. Ithaca, New York. Cornell University Library. 1933.
13. McCay, Jeanette Beyer. Responses of the individual children to the menus and to individual foods during the year. Unpublished M.S. Thesis. Ithaca, New York. Cornell University Library. 1933.
14. McCay, Jeanette B., Waring, Ethel B. and Kruse, Paul J. Learning by children at noon-meal in a nursery school: ten "good" eaters and ten "poor" eaters. Genetic Psychology Monographs. 22: 491–555. 1940.
15. Rand, Winifred, Sweeny, Mary E. and Vincent, E. Lee. Growth and Development of the Young Child. 3rd ed. p. 231. Philadelphia. W. B. Saunders Co. 1941.
16. Stuart, Harold C. Healthy Childhood. p. 197. New York. D. Appleton–Century Co., Inc. 1933.
17. Thom, Douglas A. Everyday Problems of the Everyday Child. p. 68. New York. D. Appleton and Co. 1929.
18. Wert, James E. Educational Statistics. p. 183. New York. McGraw–Hill Book Co., Inc. 1938.
19. Yule, G. Udny and Kendall, M. G. An Introduction to the Theory of Statistics. 11th ed. p. 455. London. Charles Griffin and Co. 1937.

VALUE OF CONTROLLED READING: SUMMARY OF OPINION AND RESEARCH

ARTHUR E. TRAXLER
Educational Records Bureau, New York City

Among the controversial questions relating to methods of teaching pupils to read, perhaps none has been debated more persistently and vigorously in recent years than the question of the importance and value of training in the mechanics of eye-movements. It is known that good readers have regular eye-movements and that poor readers do not. On that point, there is a wealth of incontrovertible evidence. There is, however, a difference of opinion concerning the significance of the relationship between eye-movements and reading skill. Some authorities regard the complex of psychological factors involved in reading comprehension as the all-important thing and hold that eye-movements are only symptoms of the effectiveness of comprehension. They insist that if pupils are instructed in such a way that they learn to understand what they read, their eye-movements will as a rule be satisfactorily mature.

Others believe that faulty eye-movements are frequently a cause of reading inefficiency and that noteworthy improvement in reading skill may be brought about by direct training of eye-movements. As a result of this point of view, certain instruments have been developed to "pace" reading speed. One of these is the metron-o-scope,[1] a triple-action electrically operated tachistoscope consisting of three synchronized rotary drums designed to present successive thirds of a line of reading matter. When the first line of a selection has been presented in this way, the second line appears in the same manner through the same openings, and so on until the entire selection has been read. A group of people may read from the metron-o-scope at one time. Through the use of different rolls of reading matter prepared for the metron-o-scope, a certain amount of variety can be introduced into the reading program.

The other outstanding device for use in a program of controlled reading is the Harvard Films for the Improvement of Reading.[1] Under this plan, the reading material is projected on a screen, and successive phrases or thought units of the material are lighted up. The reader's eyes follow the lighted portion of the screen, and thus the number and duration of his eye-movements are largely controlled. The films, like the metron-o-scope, can be used with a group.

The Harvard Films provide for somewhat more flexibility than the metron-o-scope, but the basic theory back of the two devices is about the same. These instruments have sharpened and brought to a focus the difference of opinion among reading specialists concerning the value of training in the mechanical aspects of the reading process. Thus, the controversy has tended to center around them —particularly about the metron-o-scope, which is the older of the two.

Since the question of eye-movement training has a definite and practical relation to the way in which teachers will give instruction in reading, it would seem that a summary of some of the significant literature on the subject should prove helpful. In this article, there is an attempt, first, to present the written statements of various authorities and, second, to summarize such research data as are available on the question.

EXPERT OPINION[1]

Several authorities on reading have indicated that devices for controlled reading, such as the metron-o-scope, have very definite values in a reading program. Center and Persons have made considerable use of the metron-o-scope in their Reading Clinic. Center (15)[2] has stated that "Another machine which we have found very useful in the Reading Clinic is the tachistoscope, an instrument which permits controlled reading.

[1] Distributed by the American Optical Company.

[1] Distributed by the Harvard Psycho-Education Clinic. Cambridge, Massachusetts.

[1] Many hours were spent in combing the literature in an attempt to find statements that would accurately reflect points of view. It is realized, however, that the quotations in this section may not represent either the best or the most recent published statements of the various individuals quoted. Some outstanding reading specialists are not included because no statements by them on this particular topic are at hand.

[2] The numbers in the parentheses refer to the bibliography at the end of the article.

The model that we use is the metron-o-scope, made by the American Optical Company. The metron-o-scope is a three-shutter device that trains the pupil in regular, rhythmic, left-to-right eye-movements, and is so constructed that the speed of reading can be adjusted to the reader and gradually increased as his improvement warrants. A pupil's comprehension of what he reads is checked, for speed without comprehension is of no value in reading."

Betts (5), who has also used the metron-o-scope in his Reading Clinic, pointed out that "Instruments such as the metron-o-scope are frequently called 'pacing' devices. The implication is that by a direct mechanical pacing of the eye-movements, rhythmical and efficient reading habits will be established. Although the value of such instruments is not questioned, the typical explanation of what is actually happening needs scientific appraisal. Three situations are created: 1) The individual is given practice in reading materials printed in large-size type. 2) Habits of rapid perception are established. The reader soon learns that regressive eye-movements take time and that the initial perception must be accurate as well as rapid. 3) The work is well motivated, the goal being to get the meaning of the phrase during the short exposure period. In certain situations, these three items alone probably are sufficient to bring about improvement in reading ability which is reflected in more efficient eye-movements. The point is that eye-movements *are* indirectly improved. The metron-o-scope is an effective device for developing rhythmical reading and rapid perception. It should be a part of the equipment of a reading clinic."

The idea that eye-movements are a causative factor in reading ability was stated succinctly by Cole (17) as follows: "The good reader is good because he uses few fixations, rarely makes a regression, and hits the beginning of each line accurately." Cole also said that "Good mechanics automatically lay the foundation for good comprehension and permit a child to understand what he reads as well as he can, in view of his intelligence and experience."

Stone (57) stated that "Eye-movement training using reading material with extra spacing between word groups is an aid in increasing rate in straight-ahead reading." After summarizing and commenting favorably on an experiment by Robinson in improving the eye-movements of college freshmen, Stone continued "An instrument called the metron-o-scope developed by Mr. J. Y. Taylor, and distributed by the American Optical Company, is a device to control mechanically the exposure of reading material for developing rhythmical eye-movements, and increased span of recognition, and consequently a more rapid rate of reading."

Some reading specialists have accepted improvement in eye-movement habits as a legitimate objective of training in reading but have changed the emphasis on the training process itself. McCallister (43), for example, has said that "Training in recognition should have as its objectives accuracy of recognition, a wide span of recognition, rhythmic progress of perceptions along the line, and accurate return sweeps from the end of one line to the beginning of the next." But he has gone ahead to point out that "Such training is probably most effective when emphasis is placed on thought getting or acquiring understandings."

In one of his summaries of research in reading, Gray (33) pointed out that "The general teaching procedures referred to have been supplemented by a series of special remedial techniques which have excited wide attention. Among them are the kinaesthetic method of promoting word recognition; the use of the flash meter in developing greater accuracy in perceiving words; the use of the metron-o-scope and films in presenting reading material under controlled conditions of word spacing and time in order to stimulate directional movements of the eyes in reading, to increase the span of recognition, and to increase speed of reading. Evidence has already been secured of the value of each of these devices. Their superiority to other methods of achieving the same ends has not been experimentally established."

Buswell's researches in eye-movement habits some twenty years ago (12) have been regarded by some workers in the field of reading as providing a factual basis for controlled reading. There is some justification for this conclusion, and Buswell himself in two recent experiments in the improvement of reading (13, 14) set up certain scientific procedures for developing regularity of eye-movements and reducing the number of re-

gressions. Buswell has, however, clearly stated the theory underlying this type of laboratory technique and has warned against purely mechanical eye-movement training. In one of his publications, he said "Because of the confusion in some of the literature in regard to enlarging the span of recognition, it is highly important that teachers understand the situation clearly. The assumption all too frequently encountered, namely, that physiological exercises in which eye-movements are produced by sweeping the eye across a series of dots or lines will increase span of recognition, misses the point entirely. Of course, reading involves peripheral adjustments through eye-movements, but training eye-movements does not increase reading ability. Neither does the correction of visual defects, in itself, improve reading. This type of work may rightly be described as 'mechanics' and has brought to the term 'mechanics' the deserved contempt which many people feel for it.

"The exploiting of machines and gadgets by persons who do not understand the psychology of reading seems at present to be adding greatly to this mechanistic folly. Because the span of recognition can be measured objectively by photographing eye-movements, it does not follow that by stimulating eye-movements mechanically a wide span of recognition in reading can be developed. Amateurish practices of this kind should not be permitted to discredit scientific techniques, which when legitimately used, possess real value. Furthermore, a clear distinction should be made between methods of teaching reading which eventuate in the development of a broad span of recognition such as normally accompanies efficient reading and methods which attempt to gain the same end by a purely mechanical control of muscular reactions. Certain forms of apparatus may be useful in stimulating a type of reading which will result in a broad span of recognition. However, apparatus can never be a substitute for educational theory; it only implements educational theory at certain points."

In the same publication, Buswell stated further that "There is no doubt that at the present time remedial reading is being grossly exploited. It is being exploited, on the one hand, by over-ambitious programs which come from within the school itself. On the other hand, it is being exploited by commercial concerns who are loading public schools with expensive apparatus which is impressive only to the uninformed and which is being emphasized out of all proportion to its actual value in a practical remedial program. Some clear thinking is needed to prevent remedial programs from out-weighing developmental programs." (14)

Strang (58) has also criticized at some length instruments for controlling eye-movements. After describing the metron-o-scope and indicating that it has high interest value and may help in breaking up vocalization, she said, "Three possible disadvantages are: 1) an individual's own grouping and phrasing may not correspond to the groupings of words presented, some of which are decidedly not natural; 2) the focusing of attention on speed may interfere with the mature reader's habits of comprehension; and 3) attention may be attracted to mechanics rather than to meaning. It is possible that practice with the metron-o-scope may not only be futile but prejudicial to some individuals.

"An improvement of the metron-o-scope technique is the use of moving-picture films in which one phrase at a time stands out clearly from a dimly perceived page. The moving-picture film is definitely superior to the metron-o-scope in its greater flexibility . . .

"In most of the experiments, a combination of methods has been used so that the improvement resulting cannot be attributed to any one mechanical device. When mechanical devices are used with high-school and college students, there may be some danger of disorganization of the effective habits which they already have built. There is also the tendency to direct attention from the thought to the mechanics and to decrease rather than increase those most important characteristics of good readers—flexibility and adaptability."

In *Better Reading Instruction*, prepared by the Research Division of the National Education Association (4), the point of view is presented that "Whenever faulty eye-movement habits are discovered, teachers should regard them as *symptoms* of some fundamental difficulty—not as *causes* of poor reading. They are to be eliminated by finding and correcting the real difficulty not by attempting

to 'pace' the eye-movements as some have attempted to do. In other words, desirable eye-movement habits are not developed by direct teaching, but are a by-product of correct reading instruction."

Sisson (55) agrees with Strang that it may not be advisable to direct the reader's attention to the mechanics of eye-movements. His statement is as follows: "In summary, it may be said that eye-movement training is no more efficacious, as regards reading rate, than mere reading with intent to improve; and it may, through directing the reader's attention to the mechanical processes of ocular movement, detract from the understanding of what is read."

The possible disadvantage of the metron-o-scope due to lack of flexibility in grouping and phrasing as mentioned by Strang has been emphasized by Harris (35), who said that "One of the difficulties with the metron-o-scope is that it assumes that good readers read at a uniform rate throughout a selection, and make exactly three fixations on each line. At least, the training it provides is designed to produce these effects. Eye-movement studies show, however, that good readers vary their rate somewhat during the reading of a selection, and that few if any of them consistently make three fixations to each line. Instead their rate and fixations continually change according to the material they read.

"Dearborn has described some experiments in the use of motion-picture apparatus for presenting reading material phrase by phrase at a controlled rate of speed. A full page is thrown on the screen at one time, and the speed of reading is controlled by the fact that one phrase at a time stands out distinctly against the rest of the page. Speed can be changed by varying the rate at which the film unwinds. This technique seems to approximate a normal reading situation more closely than the metron-o-scope does . . ."

Gates (24), like Buswell, has urged the need for expert knowledge if special devices for remedial training are to be used. In one of his writings, he said that "Complex instruments, of course, are useful only if the investigator understands them and uses them expertly. No devices or machines automatically diagnose or remedy reading defects."

A point of view closely similar to that expressed in *Better Reading Instruction* is set forth by Bond and Bond (9) as follows: "There are two contradictory points of view regarding the role of eye-movements in reading. It is the unquestioned fact that good readers of any given material have effective eye-movements. Some research workers in the field of reading believe that faulty eye-movements are the *cause* of reading disability. Others believe that they are the *result* of inability to read the material adequately. The writers hold the latter point of view. The problem does not seem to be one of correcting faulty eye-movements; rather it is correcting the poor reading techniques, of which faulty eye-movements are simply a result."

Tinker (60), who over a period of many years has contributed numerous eye-movement studies to the literature on reading, stated in one of his articles that there is lack of evidence that training eye-movements, as such, develops effective habits which improve reading ability, and held that eye-movement patterns do not cause, but merely reflect, efficient or poor reading performance.

Certain harmful effects that might result from the use of the metron-o-scope were pointed out by Blair (7). He advocated a simple push-card method for the development of speed in reading. He said that "The push-card method not only provides the pushing feature, but also utilizes the school's regular reading materials and allows the pupil to read ahead as fast as he desires or to drop back. The push-card device thus provides a method of increasing speed while retaining many of the elements of the natural reading process."

The point of view of a supervisor of a remedial project in a large city was expressed by Scarlet (50), who said that "Research studies have not yet substantiated the claims made for the machines. Rhythm reading, reading for speed—these are but nostrums hawked about in the educational marketplace. *Caveat emptor.* Salvation does not lie in mechanical treatment." In a recent article, Pilant (47) was even more outspoken in criticism of this type of training. He said that ". . . programs which set out merely to raise the reading rate of a certain percentage, or to increase the vocabulary by so many words, or to raise reading mileage, or to establish non-regressive eye-movements—start eyes from the left end of the line, attain rhythmic sweep from line to line—or to rec-

ognize topic sentences, are failures to start with. Reading worthy of the name is more than speed, eye-movements, muscle exercise, optical correction—it is understanding, a way of learning what one wants to know.

". . . Machines cannot teach those to read who see no sense or pleasure in reading; machines are not necessary for those who do. Those who have a burning desire to read, to read rapidly, to read creatively, will do so. Of course, that does not mean that the way to mastery cannot be smoothed for them; but it does mean that narrow mechanical reading objectives, apparatus, and habits may merely distract them from the only genuine purpose in reading, which is the better understanding of one's self and society. Machines appeal to poor readers because they are something new, strange, mysterious. This awe of the machine does help create a proper mind-set for breaking with old habits and putting on new virtue. But it also tends to mislead the mediocre into believing that the machine will do what the mind cannot or the will is too lazy to attempt."

Some of the critics of the devices for giving controlled training in reading have perhaps neglected the fact that these instruments do not prevent the teacher from emphasizing understanding and are not necessarily in conflict with a reading program that is designed to improve comprehension. Dearborn (18), who together with Anderson developed the Harvard Films, stresses understanding and indicates that the motion-picture technique is designed not to stretch the visual span but to increase the perceptual span and improve reading habits through bringing fresh motivation to the reader.

The importance of motivation in controlled reading techniques has also been recognized by Wilking (65), who has set forth another point which is worth consideration: "It has been my observation with college students that a great majority of them do not read rapidly because they are afraid to—that is, if they attempt to speed up, they immediately feel that comprehension is going by the board. Through the simple expedient of having them miss all if they lag behind, the films compel them to progress at a respectable rate at which they soon discover that they are still able to comprehend. Once this is realized, they are able to transfer this newfound speed to their regular reading matter. The films have an additional advantage of being intrinsic motivation. Their novelty holds attention and they have all the blandishments of the mechanical or scientific. For these reasons, the film method of improving reading speed, regardless of one's biases or convictions as regards the feasibility of training eye-movements, can be very successful."

The motivating value of devices for controlling reading has also been pointed to by Durrell and by Bear. As stated by Durrell (21), "Nothing is more important in an instructional program in reading than that every lesson—every exercise—be so motivated that interest and attention will be maintained at a high level. The use of such devices as the metron-o-scope, tachistoscope, and motion-picture and lantern slide projectors will help to vary the classroom instruction." Bear (3), in referring to the use of the metron-o-scope in the remedial reading program for freshmen at Dartmouth, said: "It has value in motivation and some students (usually those who read slightly slower than the machine is operating at a given time) feel that they do book reading with more ease immediately following practice on the machine."

Another possible value in the use of these devices has been pointed to by Louise Farwell Davis (29) in the following words: "If most of us use our eyes more effectively at a far point, why don't we present more material in the metron-o-scope or similar instruments . . . Why do we depend almost entirely on a purely visual approach at the near point in teaching children to read or spell when examiners other than ourselves discover the visual mechanism to be inadequate for the constant use which school activities require?"

The foregoing paragraphs do not exhaust the expressions to be found in the literature concerning the desirability of controlling eye-movements, but they are sufficient to show that specialists vary widely in their point of view concerning this question. Let us now examine the evidence provided by the research literature dealing with this subject.

Research Studies

The situation with respect to the evaluation of procedures in teaching reading has been characterized by Strang (30) in the fol-

lowing words: "It is difficult to evaluate the relative effectiveness of different methods and devices because there have been few controlled experiments with high-school and college students in which any of these devices has been the sole means of training. For example, spaced reading has been used at the same time as free reading; tracing has been combined with tachistoscopic drill and written summaries. Usually the device has been one of a number of methods used, and the improvement made might have been due to the wide reading, the word study, or other methods which were used simultaneously with the metron-o-scope or pacing drills. Moreover, any new method is usually presented with enthusiasm and optimism by the investigator, and the subject is thereby stimulated and encouraged to use his latent energy. This element of faith in a device may be one factor in the success reported in the case of a number of remedial measures. Of course, any device should be used as part of the natural reading process as a whole, not in isolation. Thus used, it is more likely to function in a student's everyday reading."

One of the first experiments with the metron-o-scope was reported by Taylor (59), the inventor of the machine. He set up experimental and control groups, each consisting of twenty-five high-school pupils. Each group was given practice in rapid reading during thirty-minute practice periods on ten consecutive school days. The control group read short stories from mimeographed sheets while the experimental group read the same material from the metron-o-scope. Photographs of eye-movements in a test of 148 words administered before and after training showed that the reduction in number of fixations and in regressive movements was significantly greater for the experimental group than for the control group. Neither the experimental nor the control group made significant improvement in comprehension, as measured by the tests given before and after training. The importance of the latter point was emphasized by Witty and Kopel (67), who criticized Taylor's study. They stated that one must conclude that metron-o-scope reading may effect change in the mechanical process but apparently does not influence the process of meaningful reading.

Center and Persons (16) also began experimenting with the metro-o-scope soon after this instrument became available. They described an experiment in remedial training with the metron-o-scope with eight pupils who had very low comprehension scores. Their rate scores were approximately equal. These pupils were given eight training periods of twenty-five minutes or two sessions a week for a month. The purpose was to increase comprehension, develop rhythmical reading, reduce number of regressions, and increase speed by reducing the number as well as the duration of fixations. The initial rate of reading, 180 words per minute, was gradually increased by means of metron-o-scope work, and during the final periods the rate was increased to 280 words per minute with no discomfort. After each period of silent reading, comprehension tests were given. Ophthalm-o-graph reading records before and after metron-o-scope training showed improved comprehension, reduction in number of regressions and fixations, and consequent increase in the span of recognition. There was no control group in this study.

About the time the metron-o-scope was being perfected, Triggs (62) carried on a small controlled study with six freshmen at the University of Chicago. Three students were chosen for the experimental group and three for the control group. The students were matched on the basis of scores on a battery of reading tests. The three experimental subjects were trained one-half hour a day on the metron-o-scope and one-half hour a day on easy interesting reading with emphasis on speed. The control students were not trained. The experimental students made marked progress in rate as compared with the control students.

Two other instances in which the metron-o-scope has been used in an experimental way at the college level have been reported. In their Dartmouth Study of Visual Factors in Reading, Imus, Rothney, and Bear (37) reported the use of the metron-o-scope in an attempt to correct the reading habits of students who regressed frequently and to help some types of plodding readers. They stated that the metron-o-scope proved to be an interesting and stimulating variation in procedure but reported no specific results based on this kind of training alone.

Lauer, Henry, and Fritz (39) reported an investigation the purpose of which was to compare four methods attempting to improve

the reading ability of college students as follows: (1) Instruction and motivation in which students were given suggestions for improving their ability and encouraged to do so. (2) Instruction and motivation plus a home-study project in which students timed themselves while doing a portion of their regular studying and plotted their record. (3) Instruction and motivation plus group practice in reading mimeographed sheets. (4) Instruction and motivation plus group practice in reading from the metron-o-scope. Comparisons were made with a non-practice control group. The results were evaluated by means of the Iowa Silent Reading Test, the Minnesota Speed of Reading Test, and eye-movement records on the ophthalm-o-graph. Definite improvement in test scores was found by all methods of training. The number of fixations and regressions of the eyes increased in all groups except the one trained on the metron-o-scope. It is difficult to reconcile the increase in number of fixations and regressions with improvement in reading test scores.

Several public schools have used the metron-o-scope as a part of the regular remedial reading program and have tried to evaluate the results in a somewhat informal way using techniques that are not highly statistical. For example, Hallock (34) described the basic remedial reading program in the Central High School, Valley Stream, New York. In that program, pupils whose tests reveal the need for remedial reading are given special corrective work. Much of the reading training is done by means of a metron-o-scope. A daily record is kept of the pupil's work in both speed and comprehension. Photographs of a pupil's eye-movements taken before training with a metron-o-scope and after twenty weeks of such training show some improvement in mechanics accompanied by improvement in comprehension. Similarly, Dodson (20) reported the use of the metron-o-scope daily in reading classes and stated that "While we cannot make any sweeping claims for our new approach to the reading problem, we do know that the number of reading failures decreased this year and that the average number of books read by each child in each grade increased." Smith (56) set up an experiment to determine the effectiveness of metron-o-scopic instruction in the fifth and sixth grades of the Crockett School in El Paso, Texas. However, the metron-o-scope was used as only one device to supplement the existing instructional methods and procedures of teaching reading. It was stated that the entire group made more than average progress in reading, and that some children made exceptional progress. Head turning and lip reading practically disappeared by the end of the year.

A metron-o-scope training program in the South Ward Elementary School of Brownwood, Texas, was evaluated by Lee (41, 42). Records of the pupils' progress in reading were kept and compared with previous records and reading records of schools not using the metron-o-scope. Data were obtained on standardized reading tests before and after eight and a half months of metron-o-scope training. Lee gave the following results from the use of the metron-o-scope: (1) Lip reading was almost eliminated. (2) The number of regressions was reduced. (3) The duration of fixations was reduced. (4) The number of fixations was reduced. (5) Eye span was increased. (6) Speed was increased. (7) Comprehension was increased. (8) Better reading attitude was created. (9) Failures in reading and reading subjects were reduced. The techniques and conclusions of Lee's study were criticized by Moore (46). Moore criticized especially Lee's apparent faith in the comprehension test of the ophthalm-o-graph, which Broom (10) had previously shown to be unreliable.

Garver and Matthews (23) analyzed the results of speed drills with the metron-o-scope given to pupils in a slow group of a ninth grade. Twenty pairs of pupils formed an experimental group and a control group. The pupils were paired as to intelligence, reading quotient, grade placement in rate and comprehension, and sex. Both groups were tested before and after the training period with the Iowa Silent Reading Test, and their eye-movements were photographed with the ophthalm-o-graph. The experimental group had speed drills on the metron-o-scope as a part of its regular English period for ten weeks, while the control group did not. The authors reported that the experimental group increased its rate of reading nearly five times over that of the control group and finished at a rate thirty-two words per minute above the control group. Both groups gained in

comprehension over the ten-week period, but the difference in gains was insignificant. The procedures and conclusions of Garver and Matthews were criticized by Simpson (54). Simpson pointed out especially that the rate part of the Iowa test involves only two minutes while the other five tests included in the comprehension score are timed tests so that rate is a major factor. He suggested that the thirty-three minutes of timed reading in the five comprehension parts was more significant than the two-minute rate test, and asked whether the results did not indicate that the training by means of the metron-o-scope was of doubtful value.

One of the most extensive series of studies in which the metron-o-scope has been employed regularly as a part of the training program in a reading clinic has been reported by Sievers. The first study (51) dealt with the freshman reading program at the University of Wichita during the school year 1938–39. Heterogeneous groups were trained in two one-hour periods each week for nine weeks. Included in the training were drill on the metron-o-scope as a part of every lesson, reading speed tests, reading practice exercises, and drill with the metronome as part of some lessons. Complete records for thirty-eight freshmen in the trained group who were matched with thirty-eight untrained freshmen as a control group showed considerable improvement during the training period. The median rate and comprehension scores of the trained group were 6.40 and 8.57 points, respectively, below the norms at the beginning of the program. The rate and comprehension medians were 8.00 and 5.80 points above the norms at the completion of the training. Comparison with the data for the control group indicated that the increase in scores of the training group was the result of the reading program and not the three months of college experience.

In the second study, Sievers (52) described and evaluated the use of the metron-o-scope with children in the University of Wichita Reading Clinic. In answer to the question, "Does the metron-o-scope hurt the eyes of the child as he reads?" Sievers stated that ". . . all evidence indicates, both subjectively and objectively, that the rhythmic metron-o-scopic training does not hurt the eyes of the child." In the reading program of the clinic, the metron-o-scope was used only as an additional technique to aid in the process of learning to read. It was pointed out that the habits learned on the metron-o-scope may not transfer to textbook reading unless efforts are made to secure this transfer. The metronome was found to be an aid in the solution of this problem. In summarizing the results of the whole training program of which the metron-o-scope was a part, Sievers stated that the average gain per month in silent reading achievement was three months, and that in oral reading the average gain per month was 4.9 months. There was marked variation among the individual pupils in amount of gain.

The purpose of Sievers' third experiment (53) was to determine whether the general reading program in operation in the Wichita Public Schools should be supplemented by the following techniques: (1) metron-o-scopic, (2) metronomic, (3) practice reading exercises, (4) kinaesthetic aids for retarded readers. An elementary school and two intermediate schools were selected for reading diagnosis and training. The training period lasted three months. The results indicated that the use of these techniques in addition to the general reading program of the schools produced gains three times as great as were to be expected of an average group during a three-month period. It was concluded that the techniques have value in a reading program, whether it be class, corrective, or remedial in nature, and their inclusion in a general reading program was suggested and endorsed.

One of the better statistical studies of controlled reading thus far made was carried on by Witzeman (68). This study was concerned with the diagnosis and treatment of reading defects by use of the ophthalm-o-graph and the metron-o-scope of four hundred students of an Akron, Ohio, High School. The changes produced by the metron-o-scope were evaluated by means of ophthalm-o-graph records, scores on the Iowa Silent Reading Test, and scores on the Monroe Standardized Silent Reading Test. Two hundred pupils, composed largely of inadequate readers from the freshman and sophomore classes, were given thirty-six remedial lessons with the metron-o-scope. On the Iowa Silent Reading Test, they gained 1.66 grades in comprehension and 1.90 grades in rate. Their scores on the Monroe test also

showed significant improvement. In addition, twenty-five pupils in the freshman and sophomore classes were matched with twenty-five others in the same classes, and two matched groups each containing forty seniors were also set up. The experimental group in each pair was given metron-o-scope training while the control groups received no special training. The gains of the experimental groups were larger than those of the control groups.

In a recent article, Anderson (1) summarized studies by Walker, Anderson, Fairbanks, Swanson, and Young which indicated that eye-movements are not really the cause of good and poor reading. He concluded from the evidence presented that the metron-o-scope had no control over those factors which ultimately make for good comprehrension, that is, intelligence, experience, and vocabulary. He suggested that all that could be hoped for from such training is that it would succeed in teaching the individual to read in larger units and to react to words more quickly, and stated that if the eye-movements of the poor reader are made like those of the good reader, this does not mean that the two will thereafter comprehend alike. He does not deny that there is some place for the metron-o-scope in the reading program, but he would assign it about the same place that flash cards have.

Most of the research studies dealing with controlled reading have employed the metron-o-scope as the training device, but one study using the motion picture films was carried on by June Foy (22) in two elementary schools in Syracuse, New York, and reported by Troyer (63), under whose direction Miss Foy worked. Experimental and control groups each consisting of thirty-eight pupils were equated on the basis of sex, chronological age, mental age, reading speed, visual acuity, and the Gates Silent Reading Test, Form I. The experimental group received three thirty-minute practice periods per week for six weeks on stories and nature study materials projected from motion picture films in successive phrases. The median gain of the experimental group over the control group, as shown by Form II of the Gates test, was twelve months. Interestingly enough, the gains in speed were most marked among the better readers.

Dearborn and Wilking (18) have recently published the procedures and results of three experiments in improving the reading of college freshmen. In these experiments, the Harvard Films were one of several techniques used. Gains were measured by means of the Iowa Silent Reading Tests, the Nelson–Denny Reading Test, and the Cooperative Reading Test. The results indicated that the remedial groups profited significantly from the remedial sessions.

A doctor's dissertation by Harris (36) at the University of Chicago does not deal directly with devices for controlling reading, but it has important implications for them. The central problem of Harris' study was to investigate the relation of a group of ten psychological factors and a group of nine physiological or visual factors to the span of recognition. The study dealt with the relationship of these factors to the span of recognition at seven grade levels, the second through the sixth, the ninth, and the thirteenth. Laboratory techniques and standardized tests were employed in the study. The data were employed by means of correlation analysis and case-study techniques. The main conclusion of the study was that the psychological factors selected were much more highly correlated with span of recognition than were the visual factors. An important implication of the study is that if training is to be given which will develop a broad and mature span of recognition, basic attention should be given to the psychological factors of the reading process.

Summary and Implications

The results of the research studies that have been conducted thus far do not provide clear-cut evidence either favorable or unfavorable to controlled reading. The experimental procedures and the conclusions of a considerable proportion of the limited number of available studies would not meet the criteria of an acceptable statistical study. Notwithstanding the limitations in the data as a whole, however, it appears that the findings are somewhat more favorable to the value of controlled reading techniques than is the sum total of the opinion of experts in the field of reading. Most of the studies do suggest that considerable improvement is made under teaching procedures employing controlled reading. There are, however, two limitations which greatly interfere with definite conclusions in this regard. One limita-

tion is that information concerning the permanence of the gains in test scores brought about by controlled reading is almost nonexistent. The other limitation is that, as Strang has pointed out, controlled reading has usually been only one of a number of techniques employed and it is almost impossible to say just what procedures have been responsible for the improvement shown by the subject.

The logic of the argument of those who hold that comprehension is all-important and that eye-movements are a result and not a cause of reading ability seems so sound to the writer that he has difficulty in seeing any justification for purely mechanical attempts to train pupils in better eye-movement habits. It is his opinion, however, that controlled reading techniques need not be used in a purely mechanical way but that they might well be used to supplement, in a way that would have strong motivating force, a program in which the building up of reading comprehension is emphasized. It is altogether possible, as Dearborn, Wilking, and various other reading specialists have suggested, that controlled reading procedures have important values aside from the question of the correctness of direct training of eye-movements as such.

At any rate, it seems evident that controlled reading techniques should have the benefit of a more extensive, detailed, and rigidly controlled experimental study than has yet been made in this area. Two questions are particularly pertinent: 1) Can an experienced teacher of reading get better results through the use of instruments for controlling reading than he can obtain with more conventional procedures of reading instruction? 2) Can an inexperienced teacher of reading bring about as much improvement in his pupils by employing controlled reading techniques as a good experienced teacher can secure through more informal procedures carefully planned and carried out? A large-scale study of this kind would call for the cooperation of classroom teachers, psychologists, and statisticians. If the study were sufficiently comprehensive and were thoroughly defensible from experimental and statistical considerations, it would be of much practical help to teachers of reading in elementary schools, high schools, and colleges throughout the United States.

Bibliography

1. Anderson, Irving H. "The Ophthalm-O-Graph and Metron-O-Scope Evaluated in the Light of Recent Research on the Psychology of Reading," *The Teachers College Journal,* XII (January, 1941), 62–63.
2. Barry, Linda; Madden, Mable; and Pratt, Marjorie. "Reading Difficulties of High-School Pupils," *School Review,* XLVI (January, 1938), 44–47.
3. Bear, Robert M. "The Dartmouth Program for Diagnostic and Remedial Reading with Special Reference to Visual Factors," *Educational Record,* XX, Supplement No. 12 (January, 1939), 69–88.
4. "Better Reading Instruction," *National Education Association, Research Bulletin,* XIII (November, 1935). Washington, D. C.: Research Division of the National Education Association, 1935.
5. Betts, Emmett A. *The Prevention and Correction of Reading Difficulties.* Evanston, Illinois: Row, Peterson and Company, 1936.
6. Betts, Emmett A.; Donnelly, Helen E.; Aaron, Sadie; and Everett, Mabel. "Challenging the Learner," Report of the Summer Session Reading Clinic, State Normal School, Oswego, New York. *Elementary English Review,* XV (April, 1938), 149–58.
7. Blair, Glenn Myers. "Instrumentation and the Remedial Reading Program," *Educational Administration and Supervision,* XXVI (March, 1940), 205–6.
8. Blair, Glenn Myers. "Remedial-Reading Programs in Senior High Schools," *School Review,* XLIX (January, 1941), 32–41.
9. Bond, Guy L.; and Bond, Eva. *Developmental Reading in High School.* New York: The Macmillan Company, 1941.
10. Broom, M. E. "Reliability of the Reading Graph Yielded by the Ophthalm-o-graph," *School and Society,* LII (September 14, 1940), 205–8.
11. Bureau of Visual Science, American Optical Company. "The Metron-O-Scope," *The Ophthalm-O-Graph, The Metron-O-Scope, Manual for Controlled Reading,* 1938.

12. Buswell, Guy T. *Fundamental Reading Habits: A Study of Their Development.* Supplementary Educational Monographs, No. 21. Chicago: Department of Education, University of Chicago, 1922.
13. Buswell, Guy T. *How Adults Read.* Supplementary Educational Monographs, No. 45. Chicago: Department of Education, University of Chicago, 1937.
14. Buswell, Guy T. *Remedial Reading at the College and Adult Levels: An Experimental Study.* Supplementary Educational Monographs, No. 50. Chicago: Department of Education, University of Chicago, 1939.
15. Center, Stella S. "The Significance of the Reading Clinic," *The English Journal,* XXVII (May, 1938), 382.
16. Center, Stella S.; and Persons, Gladys L. *Teaching High-School Students to Read.* New York: D. Appleton–Century Company, 1937.
17. Cole, Luella W. *The Improvement of Reading with Special Reference to Remedial Instruction.* New York: Farrar and Rinehart, 1938.
18. Dearborn, Walter F.; and Wilking, S. Vincent. "Improving the Reading of College Freshmen," *School Review,* XLIX (November, 1941), 668–78.
19. Dodson, P. J. "A Modern Program of Reading Instruction in a Public School System," *Peabody Journal of Education,* XVI (November, 1938), 179.
20. Dodson, W. L. "Mechanical Aids in a Practical Reading Program," *Texas Outlook,* XXIII, No. 12 (December, 1939), 13.
21. Durrell, Donald D. *Improvement of Basic Reading Abilities.* Yonkers-on-Hudson, New York: World Book Company, 1940.
22. Foy, M. June. *A Preliminary Evaluation of Controlled Reading Exercises Projected Upon Motion Picture Films.* Thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Education, Syracuse University, 1938.
23. Garver, F. M.; and Matthews, R. D. "An analysis of the Results of Special Drills with the Metron-O-Scope to Increase Reading Rate," *The Journal of Educational Psychology,* XXX (December, 1939), 696–98.
24. Gates, Arthur I. "Diagnosis and Treatment of Extreme Cases of Reading Disability," *The Teaching of Reading: A Second Report.* The Thirty-Sixth Yearbook of the National Society for the Study of Education. Bloomington, Illinois: Public School Publishing Company, 1937.
25. Gates, Arthur I. *The Improvement of Reading.* (Revised Edition) New York: The Macmillan Company, 1935.
26. Gilbert, Luther C.; and Gilbert, Doris Wilcox. "Reading Before the Eye-Movement Camera Versus Reading Away From It," *The Elementary School Journal,* XLII (February, 1942), 443–7.
27. Gray, William S. "Problems of Reading Disability Requiring Scientific Study," *The Elementary English Review,* XII (April, 1935), 97–98.
28. Gray, William S. "The Use of Tests in Diagnosis and Remedial Work with Special Reference to Reading and Study Habits," *Report of the Thirteenth Annual Meeting of the American College Personnel Association,* 1936. Pp. 22–27.
29. Gray, William S. (Editor) *Recent Trends in Reading.* Proceedings of the Conference on Reading Held at the University of Chicago, Volume I. Supplementary Educational Monographs, No. 49. Chicago: Department of Education, University of Chicago, 1939. (Louise Farwell Davis, "Visual Difficulties and Reading Disabilities," pp. 137–38.)
30. Gray, William S. *Reading in General Education.* Washington, D. C.: American Council on Education, 1940. (Ruth Strang, "Diagnosis and Remediation," pp. 337–38.)
31. Gray, William S. (Editor) *Reading and Pupil Development.* Proceedings of the Conference on Reading Held at the University of Chicago, Volume II. Supplementary Educational Monographs, No. 51. Chicago: Department of Education, University of Chicago, 1940. (E. W. Dolch, "Evaluation of Growth in Reading," p. 296.)
32. Gray, William S. (Editor) *Adjusting Reading Programs to Individuals.* Proceedings of the Conference on Reading Held at the University of Chicago, Volume III. Supplementary Educational

Monographs, No. 52. Chicago: Department of Education, University of Chicago, 1941. (Mildred E. Sherwood, "The Non-Reader," p. 305.)

33. Gray, William S. "Reading," *Encyclopedia of Educational Research,* p. 923. (Edited by Walter S. Monroe) New York: The Macmillan Company, 1941.
34. Hallock, Evelyn. "Diagnostic and Remedial Work in a Reading Clinic," *Nation's Schools,* XXVIII (October, 1941), 28–29.
35. Harris, Albert J. *How to Increase Reading Ability.* New York: Longmans, Green and Company, 1940.
36. Harris, Theodore Lester. *A Laboratory Study of the Relation of Selected Factors to the Span of Recognition in Silent Reading.* (Doctoral Thesis) Private Edition, distributed by the University of Chicago Libraries, Chicago, Illinois, 1941.
37. Imus, Henry A.; Rothney, John W. M.; and Bear, Robert M. *An Evaluation of Visual Factors in Reading.* Hanover, N. H.: Dartmouth College Publications, 1938.
38. Kottmeyer, William. "On the Theory of Controlling Eye-Movements in Reading," *School and Society,* LII (November 16, 1940), 496–98.
39. Lauer, Alvah R.; Henry, L. K.; and Fritz, J. F. "Improving the Reading Ability of College Students," *Psychological Bulletin,* XXXV (1938), 670.
40. Lawson, Elizabeth K. "Reading Comprehension Among College Students," *Journal of Experimental Education,* VIII (September, 1939), 103–09.
41. Lee, R. B. "Value of Metron-O-Scope Reading," *The National Elementary Principal,* XVIII (February, 1939), 109–11.
42. Lee, R. B. "Metron-O-Scope Reading," *The Peabody Reflector and Alumni News,* XII (July, 1939), 248–49; 254–55.
43. McCallister, James M. *Remedial and Corrective Instruction in Reading.* New York: D. Appleton–Century Company, 1936.
44. McNelly, A. E. *Remedial Reading at John Jay High School, Cleveland, Ohio.* (Printed through courtesy of American Optical Company.)
45. Monroe, Marion; and Backus, Bertie. *Remedial Reading.* Boston: Houghton Mifflin Company, 1937.
46. Moore, Joseph E. "Metron-O-Scope Reading: A Critical Evaluation of Mr. Lee's Study," *The Peabody Reflector and Alumni News,* XII (August, 1939), 285–86; 294.
47. Pilant, Richard. "Reading Improvement Programs," *The English Journal,* XXXI (March, 1942), 235–37.
48. Robinson, Francis P. "An Aid for Improving Reading Rate," *Journal of Educational Research,* XXVII (February, 1934), 453–55.
49. Robinson, Helen M. "Treatment of Severe Cases of Reading Disability," *Journal of Educational Research,* XXXII (March, 1939), 531–35.
50. Scarlet, Will. "Remedial Reading in the New York City High Schools," *High Points* (December 20, 1938), 18–24.
51. Sievers, Clement H. "The University of Wichita Freshman Reading Program 1937–1939," *University of Wichita Bulletin,* XIV (September, 1939), University Studies Bulletin No. 6.
52. Sievers, Clement H. "Study II: The University of Wichita Reading Clinic Diagnostic and Remediation Program," *University of Wichita Bulletin,* XIV (November, 1939), University Studies Bulletin No. 7.
53. Sievers, Clement H. "Study III: Wichita Public School Reading Program," *University of Wichita Bulletin,* XIV (November, 1939), University Studies Bulletin No. 7.
54. Simpson, Ray H. "Further Analysis of the Results of Speed Drills with the Metron-O-Scope," *Journal of Educational Psychology,* XXXII (May, 1941), 386–88.
55. Sisson, E. Donald. "Eye-Movement Training as a Means of Improving Reading Ability," *Journal of Educational Research,* XXXII (1938), 40.
56. Smith, Marguerite. "Metronoscope in Upper Grade Instruction," *Texas Outlook,* XXIV (March, 1940), 13–14.
57. Stone, Clarence R. *Better Advanced Reading.* St. Louis, Missouri: Webster Publishing Company, 1937.

58. Strang, Ruth. *Problems in the Improvement of Reading in High School and College.* (Revised Edition) Lancaster, Pennsylvania: The Science Press Printing Company, 1940.
59. Taylor, Earl A. *Controlled Reading: A Correlation of Diagnostic Teaching and Corrective Techniques.* Chicago: University of Chicago Press, 1937.
60. Tinker, Miles A. "The Role of Eye-Movements in Diagnostic and Remedial Reading," *School and Society,* XXXIX (February, 1934), 147–48.
61. Traxler, Arthur E. *Ten Years of Research in Reading.* Educational Records Bulletin No. 32. New York: Educational Records Bureau, 1941.
62. Triggs, Frances. "Remedial Reading Experiment—Rate." (Unpublished Term Paper, University of Chicago, 1936.)
63. Troyer, Muriel E. "An Experiment in Controlled Remedial Reading Using Film Materials," *Psychological Bulletin,* XXXV (1938), 669.
64. Weber, C. O. "Remedial Reading with College Freshmen," *Journal of Higher Education,* VIII (1938), 453–54.
65. Wilking, S. Vincent. "The Improvement of Reading Ability in College," *Journal of the American Association of Collegiate Registrars,* XVII (January, 1942), 183–84.
66. Witty, Paul A. "Practices in Corrective Reading in Colleges and Universities," *School and Society,* LII (November 30, 1940), 564–68.
67. Witty, Paul; and Kopel, David. *Reading and the Educative Process.* New York: Ginn and Company, 1939.
68. Witzeman, B. Evangeline. "An Experimental Study Using the Ophthalm-O-Graph and Metron-O-Scope in the Diagnosis and Treatment of Reading Defects," Department of Psychology, Western Reserve University. Published as a separate and in *The Journal of Psychology,* Vol. 11–12, (1941), 307–34.

THE SEMANTIC VARIATIONS OF CERTAIN HIGH FREQUENCY WORDS IN THE WRITTEN COMPOSITIONS OF EIGHTH-GRADE PUPILS*

HARRY K. FOSTER
University of Akron

I. PURPOSE

The purpose of this investigation was three-fold: (a) to determine the socially important meanings with which 100 selected words of high frequency of usage were used by eighth-grade pupils in a wide and representative sampling of their written compositions, (b) to determine what proportion of the total number of dictionary meanings for each of the selected 100 words eighth-grade pupils used in their written compositions, and (c) to develop a technique for handling word-meaning data on Hollerith equipment.

II. SOURCE OF DATA

1. THE SELECTION OF THE WORD LIST[1]

To insure the selection of check words showing high frequency of usage three criteria were established:

1. The word must occur within the first thousand words of Horn's (6) list.[2]

2. The word must occur within the first thousand words of Thorndike's (16) list.

3. The word must appear within at least seven of the eight lists developed by Jones (8), Wilmarth (20), Tidyman (17), Smith (2), Bauer (1), McKee (10), Fitzgerald (5), and Oberman (11) from children's writing.

From the words meeting the above criteria, 100 words were arbitrarily selected such that each would fall into one of four equal groups according to the number of meanings given it in the second unabridged edition of Webster's New International Dictionary. Words having more than 45 meanings were placed in Group I; those having 27 to 44 meanings were placed in Group II; those having 15 to 26 meanings were placed in Group III; and those having less than 15 meanings were placed in Group IV. (In counting word meanings, all meanings designated by letters and Arabic numbers in Webster's dictionary were counted except those marked obsolete.) The words meeting the criteria for frequency of usage which were selected for study in this investigation are presented in Table 1 in grouped form with the number of dictionary meanings listed for each.

2. THE PUPILS AND THE COMPOSITIONS

The check words as used in 1976 compositions written on 500 different topics during the school year 1938–1939 by eighth-grade pupils located in 38 different school systems in 13 states constituted the data for the investigation. The compositions were all written by pupils whose intelligence quotient was 90 or above.

III. HANDLING THE DATA

1. MARKING THE CHECK WORDS

The 1976 compositions, constituting the source of the check words, were read through and all words having the same spelling as the words in the check list were underlined. In underlining the words, the inflected forms of the check words, and the check words as they constituted an element in hyphenated or solid forms of compound words were omitted.

2. THE USE OF HOLLERITH EQUIPMENT

Each occurrence of the check words in the compositions was recorded on Hollerith cards together with the part of speech with which the word was used, the sentence in which the word was used, and other pertinent information concerning the location and use of the

* An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, in the Department of Education, in the Graduate College of the State University of Iowa, August, 1939.

[1] The list of 100 words selected for study was designated as the check list while the words were designated as the check words.

[2] Through an error in transcription two words, *fish* and *surprise*, occur in the check list although they are outside the first thousand words of Horn's list; *fish* occurs in the first half of the third thousand of Horn's list while *surprise* occurs in the first half of the second thousand.

TABLE 1

THE CHECK WORDS ARRANGED IN GROUPS ACCORDING TO THE NUMBER OF DICTIONARY MEANINGS AVAILABLE FOR EACH

Group I		Group II		Group III		Group IV	
Word	Number of Meanings	Word	Number of Meanings	Word	Number of Meanings	Word	Number of Meanings
back	73	bed	41	account	21	act	13
bank	45	board	29	book	26	boy	10
box	52	case	39	church	20	brother	10
call	59	cover	37	company	15	cause	10
carry	47	dress	40	country	15	child	12
clear	62	fair	37	family	17	city	9
cut	108	front	42	farm	15	cost	10
face	63	land	27	father	24	fear	10
fall	70	man	33	fish	23	friend	9
hold	50	mind	38	last	23	letter	13
light	65	number	29	mother	26	market	9
line	91	place	35	name	18	money	8
open	63	plant	34	night	16	morning	6
order	52	play	39	paper	26	need	9
point	110	reach	28	person	16	new	10
return	54	rest	41	picture	21	party	13
run	157	school	37	size	22	plan	13
short	57	send	35	store	18	please	7
side	66	sick	28	suit	16	pretty	10
spring	67	small	36	talk	20	price	9
stop	72	strong	36	town	21	rain	11
take	106	time	44	wish	15	sister	10
turn	130	train	43	woman	20	sure	9
water	57	walk	39	word	18	surprise	9
work	50	way	30	young	20	winter	9
Range	45–157		27–44		15–26		6–13

word. The words were then sorted alphabetically by machine and the various parts of speech with which each was used were grouped and totalled. The information contained on the cards was printed on large sheets which served as the working material for assigning meanings.

3. AUTHORITY FOR WORD MEANINGS

The authority accepted for defining word meanings was the second unabridged edition of Webster's New International Dictionary. This dictionary rather than some other was selected because: (a) as one of the oldest dictionaries published primarily for use in the United States it has stood the test of time for the accuracy and clarity of its definitions, (b) the second unabridged edition published in 1938, represented a later edition than did any of the other dictionaries considered, (c) the persons who contributed to this second edition are generally considered to be among the most capable in their respective fields, and (d) the general arrangement for making and defining vocabulary entries appeared to be well adapted to the purposes of the investigation. One authority rather than several was accepted because it was believed greater objectivity would thereby be obtained.

4. ASSIGNING WORD MEANINGS

Meanings were assigned independently by two different persons to the words as used in sentence context. The writer and an assistant assigned meanings to approximately one half of the words, while two assistants assigned the meanings to the other half of the words.

The two assistants who helped in assigning meanings were selected in the belief that they were above average in cultural and intellectual development and because each had received more than average training in the general field of language.

The steps in the process of assigning meanings were: (a) the sentence containing the word whose meaning was to be assigned was carefully read, (b) any combination of the check word with other words which might constitute a phrase used idiomatically was noted, (c) to determine whether or not such phrases were shown to have idiomatic meanings,[3] reference was made to the dictionary,

[3] In a few cases, when the majority judgment of the three persons assigning meanings so indicated, phrases were considered to have idiomatic meanings even though they were not listed in the dictionary.

(d) if the check word was found not to occur in a phrase having an idiomatic meaning, the definitions of the word were studied until it was felt that the one best describing its meaning was discovered, (e) if the word was found to occur in an idiomatic phrase it was so indicated and omitted from the final count of the different meanings with which the word was used. (These idiomatic usages were omitted from the final count for it was considered that the phrase and not the word constituted the meaning unit), (f) whenever two workers differed in the meaning assigned a given word, the judgment of the third person was utilized to settle the difference, (g) all meanings assigned the words were coded according to their corresponding dictionary designation.

If the sentence in which the word was used was inadequate for determining the meaning intended by the writer, reference was made to the composition from which the word was selected. For a sample of 6176 occurrences of the check words, a record of the extent to which it was necessary to refer back to the compositions was kept for each person assigning meanings to the words. For this sample, one person went back to the compositions 164 times while the second person referred back 159 times. This was interpreted to mean that sentences or parts of sentences can be made to show word meanings in a manner that is sufficiently clear so that a person assigning meanings to those words would require a more adequate context less than three times in 100.

IV. RESULTS

In order to determine which of the meanings were socially important two different bases were used; the first considered *meanings* as the unit while the second considered the *word* as the unit. In both cases the burden of usage borne by the meanings was employed as the index of social importance. When *meanings* were taken as the unit, it was arbitrarily decided, for the purposes of the study, to consider only those meanings occurring six or more times to be of sufficient social importance to subject them to further analysis; when the *word* was taken as the unit, it was arbitrarily decided that only those meanings bearing five per cent or more of the total usage of the word would be considered of sufficient importance to subject them to further critical treatment. These arbitrary decisions were necessary because no other objective measure of the social importance of meanings was available and some elimination of meanings of lower social frequency seemed necessary and desirable.

The more important findings of this investigation may be summarized as follows:

1. The check words were used in the compositions 11,030 different times with a total of 637 different meanings.

2. When the check words were considered by groups, it was found that, in general, those possessing a large number of meanings tended to be used more often and with more different meanings than did those possessing a smaller number of meanings. It was found, however, that individual words might be selected to show the reverse of this tendency.

3. When the check words were considered by groups, it was found that the median word of each group showed a gradual decrease in the number of meanings with which it was used in passing from Group I to Group IV. When the frequency of occurrence of these median words was considered, however, a similar gradual decrease was not found for the frequencies in passing from Group I to Group IV. These data were interpreted as indicating that the number of meanings with which one of the check words may be used is more dependent upon the number of meanings which the word possesses than is the frequency with which the word may occur.

4. Regardless of the criterion of social importance employed, approximately one half the total meanings with which the check words were used were found to be socially important, as here defined.

5. These socially important meanings were found to bear more than 90 per cent of the total usage of the check words.

6. No systematic relationship between the number of meanings with which the individual words were used and the number of meanings found to be socially important was noted.

7. The socially important meanings of the large majority of individual words were found to bear 70 per cent or more of the total usage of the word.

8. The average number of socially important meanings with which the grouped words were used was found to decrease in passing from Group I to Group IV. Although exceptions were noted, this was interpreted to

mean that, in general, words possessing a large number of meanings tend to be used with a larger number of socially important meanings than do words possessing a smaller number of meanings. Perhaps this is logical and to be expected.

9. The total number of dictionary meanings for all check words was found to be 3,460.

10. When *meanings* constituted the base for determining social importance, it was found that the total number of socially important meanings with which the check words were used represented 8.5 per cent of the total number of dictionary meanings.

11. When the *word* constituted the base for determining social importance, it was found that the total number of socially important meanings with which the check words were used represented 9.3 per cent of the total number of dictionary meanings.

12. When the individual words were considered, it was found that approximately one-fifth or less of the available meanings for each word are made to carry the burden of usage for the large majority of the individual words.

13. Some evidence was found which indicated that a larger per cent of available dictionary meanings are utilized for words having a small number of dictionary meanings than for words having a larger number of dictionary meanings.

V. CONCLUSIONS

On the basis of the above findings and on the basis of the experience obtained in working with these data the following conclusions appear valid:

1. For the words studied in this investigation and for the criteria of social importance here employed, approximately one half the total number of meanings with which the words are used by eighth-grade pupils are socially important.

2. For the words studied in this investigation and for the criteria of social importance here employed, approximately nine per cent of the available dictionary meanings are utilized by the socially important meanings of the words.

3. The experience gained from the study indicates that a considerable saving in time and labor can be made in word-meaning studies of this type through the use of Hollerith equipment.

VI. LIMITATIONS

In evaluating the outcomes of this investigation certain limitations should be kept in mind. First of all, the time and expense involved made it necessary to limit the investigation with respect to (a) the number of words studied, (b) the number of grades studied, and (c) the number of different types of writing used as the source of data. To generalize on the basis of the findings of the study for any one of these factors would be indeed hazardous.

The investigation is limited secondly by the subjectivity involved in assigning meanings and in recognizing idiomatic usages of the words. While this limitation could not be completely overcome, it is believed that error due to this cause was reduced to a minimum by having each meaning assigned represent the majority judgment of three persons well trained in the field of language.

The utilization of sentence context for showing word meanings constitutes a third limitation, for it is conceivable that a meaning assigned a word in sentence context may differ from the meaning assigned that word in the context of a composition. It is doubtful, however, that this limitation seriously affects the results of the study because care was taken to present each word in the sentence in which it was used by the pupil and because reference was made to the composition if any doubt existed as to the intended meaning of the word.

The study is finally limited by the fact that the word meanings were treated without relation to the part of speech with which the words were used. While it would have been possible to analyze the data on this basis it was believed that such an analysis would have clouded the major issues of the investigation.

Bibliography

1. Bauer, Nicholas, *The New Orleans Public School Spelling List*, F. F. Hansell and Brothers, New Orleans, 1916.
2. Cook, W. A., and O'Shea, M. V., *The Child and His Spelling*, The Bobbs–Merrill Company, Indianapolis, 1914.

3. Dolch, Edward W., *Reading and Word Meanings,* Ginn and Company, Chicago, 1927.
4. Fennell, Ethel L., "Recurring Words and Their Relation to Difficulties in Comprehension", *Elementary School Journal,* Vol. 29, September, 1928, pp. 42–53.
5. Fitzgerald, James A., "The Vocabulary, Spelling Errors and Situations of Fourth, Fifth, and Sixth Grade Children's Letters Written in Life Outside the School", Unpublished Doctor's Dissertation, State University of Iowa, Iowa City, 1931. (An abstract of this study is found in the *State University of Iowa Studies in Education,* Vol. 9.)
6. Horn, Ernest, *A Basic Writing Vocabulary, 10,000 Words Most Commonly Used in Writing,* College of Education, State University of Iowa, Iowa City, 1926.
7. Irmina, Sister M., "An Evaluation of the Vocabulary Content of Twelve Series of Primary Readers", The Catholic University of America, *Educational Research Bulletins,* Vol. 4, No. 7, September, 1929.
8. Jones, W. Franklin, *Concrete Investigation of the Material of English Spelling,* University of South Dakota, Vermillion, South Dakota, 1913.
9. Lorge, Irving, "The English Semantic Count", *The Teachers College Record,* Vol. 30, October, 1937, pp. 65–77.
10. McKee, Grace M., "The Vocabulary of Children's Themes", Unpublished Master's Thesis, State University of Iowa, Iowa City, 1924.
11. Oberman, Margaret H., "A Study of the Degree to Which the Written Vocabulary of Children Has Been Determined", Unpublished Master's Thesis, State University of Iowa, Iowa City, 1933.
12. Rinsland, Henry D., *The Vocabulary of Elementary School Children of the United States,* Mimeographed Report prepared by the University of Oklahoma, Norman, Oklahoma, 1938.
13. *Seventh Annual Research Bulletin of the National Conference on Research in English,* "Vocabulary Problems in the Elementary School", Scott, Foresman and Company, Chicago, 1939.
14. Studley, C. K., and Ware, A., *Common Essentials in Spelling,* Chico State Normal School, Bulletin No. 7, 1914.
15. *The Elementary English Review,* "Vocabulary Problems in the Elementary School", April–May, 1939.
16. Thorndike, Edward L., *The Teachers Word Book,* Teachers College, Columbia University, 1921.
17. Tidyman, W. R., *Survey of the Writing Vocabularies of Public School Children in Connecticut,* Teachers Leaflet, No. 15, U. S. Bureau of Education, Washington, D. C., 1915.
18. *Webster's New International Dictionary,* Second Edition, Unabridged, G. and C. Merriam Company, Springfield, Massachusetts, 1938.
19. White, Carl R., "Semantic Variations in Oral and Written Vocabularies", Unpublished Master's Thesis, State University of Iowa, Iowa City, 1933.
20. Wilmarth, Alta L., "The Vocabulary of Rural Children's Themes", Unpublished Master's Thesis, State University of Iowa, Iowa City, 1926.

AN EXPERIMENT WITH THE WEBSTER COMPLETE SYMBOLIZATION

RALPH W. HOUSE*
Pikeville College, Pikeville, Kentucky

THE PROBLEM

The anticipated difficult words in an elementary-school textbook for the intermediate or upper grades are usually placed in the glossary. In a glossary, a phonetic respelling with diacritical marks is printed on the right of each anticipated difficult word. The phonetic respelling with diacritics should be used by pupils as a phonetic aid in making an independent analysis of the difficult words which they cannot pronounce from memory when they see them printed in an incomplete symbolization.

The problem in the present study was to determine how successfully teachers who are not phonetically trained can help pupils in Grades V–VIII to master the use of a complete symbolization, as measured by an independent analysis of unfamiliar words, following a controlled experience with a specific form of instructional material.

THE EXPERIMENT

Selection of Pupils.—This study was conducted in the Training School at Pikeville College, Pikeville, Kentucky. The criteria used in the selection of the pupils were as follows: (1) the pupils must be free from foreign accent; (2) the pupils must be free from speech defects; and (3) the pupils must be making satisfactory progress in the experiencing of their curriculums. The fifty pupils used in this experiment were distributed through the grades as follows: (1) Grade VIII eleven pupils; (2) Grade VII fifteen pupils; (3) Grade VI fifteen pupils; and (4) Grade V nine pupils.

Measurement of Growth.—Pupil growth in the use of the pronunciation skills was measured by administering a test, consisting of twenty-five nonsense words. The test contained the Webster symbols for each speech sound in the English language. The test is an individual test. The test was called The Word-Analysis Test, Form A. The Word-Analysis Test, Form A was administered at the beginning of this experiment and also at the close of this experiment. The test is presented here:

THE WORD-ANALYSIS TEST, FORM A

1. yăth	9. hŭ măth′	17. săj′ŭ lăn
2. zōth	10. poi kā′ ˌ	18. sărg′ə lĭ
3. hwēngk	11. zēr dĭf′ĭn	19. kwēr ōō′mə
4. shoung	12. mā′tə wô	20. sŭf′ō nĭ
5. bōō shān′	13. shā′vĭ lĭ	21. bū shô′nē
6. dô′zŭm	14. zā lĭd′ĭj	22. īz′ē wĭk
7. nŏf′fōōl	15. nĕch′ə nĭ	23. nō tā′zhŭn
8. gēr tĕl′	16. rôzh ĕ nā′	24. där rĭzh′ŭn
		25. dī kŏt′ŭm

Only the first trial is indicative of a pupil's ability to master a complete symbolization as a phonetic aid in making an independent analysis of an unfamiliar word. If, on the first trial a pupil analyzed all of a word correctly except one syllable, the examiner would say, "You analyzed all of the word correctly except the second syllable. You missed the vowel in the second syllable. Try the word again, please. Can you pronounce it correctly this time?" If he missed the word on the second trial, the examiner would show the pupil where his analysis was wrong and ask that he try it again.

Method of Instruction.—The organic approach accompanied by an acoustic approach was used in teaching each English speech sound. The organic approach consisted of (1) a study of a number of drawings showing the most characteristic phase of the position and movement of the speech organs in articulating an English speech sound, and (2) the reading of a carefully worded paragraph describing the most characteristic phase of the position and movement of the speech organs in articulating that sound. The acoustic approach consisted of sounding key words in order that the pupils could hear the speech sound, as well as, feel the movements of the

* Assisted by the following critic teachers: Elizabeth North, and Ruth Harman.

speech organs in the production of the speech sound.

The method of instruction consisted of an audio-visual attack on the pronunciation skills. If the pupils did not know the ā skill, the teacher would help them to acquire this ā skill through the use of a known word. In helping the pupils to acquire this ā skill, the teacher used words which the pupils knew how to pronounce, e.g., sāme. The teacher's first instructional job was to give her pupils an awareness for the ā skill in the word same. This she would try to do by blending the speech sounds which are represented by the following symbols: sāme, sām, ām, and ā. The pupils were told to follow closely, and to think through the process as the teacher blended the sounds represented by these symbols: sāme, sām, ām, and ā. The teacher asked the pupils to blend outloud the speech sounds represented by sāme, sām, ām, and ā. If the instruction had ceased at this point the pupils' awareness of the ā skill would be questionable. Consequently, the teacher would write on the blackboard a one-syllable nonsense word which had the ā skill in it. The nonsense word pāmz was used. The teacher blended the speech sounds represented by the nonsense word pāmz: pāmz, āmz, ām, and ā. She then asked one or more of her pupils to blend the speech sounds represented by the symbols in the word pāmz. The pupils practiced blending the speech sounds represented by the nonsense words on that page in the program of instructional material which introduced the ā skill.

The word-whole method of analyzing words was emphasized. But, if a pupil was unable to look at the nonsense word in the square spāg and pronounce it correctly, he was encouraged to use an analytic method of analyzing the word. In the analytic method the pupil was told to look at the word spāg in his program of instructional material. The word spāg was printed as follows:

ā
āg
pāg
spāg

The pupil was told to look at the ā and sound out the word spāg, viz., ā, āg, pāg, and spāg.

The development of a desire by the pupils to master the pronunciation skills was aroused by spirited contests or games. These contests resembled football games, baseball games, and the old-fashioned spelling matches.

Instructional Material.—A program of instructional material was prepared for use in this experiment. A page in the program of instructional material introduced the Webster complete symbolization for only one English speech sound, e.g., ā as in ate. The organization of each page was as follows: (1) at the top of the page was a number of drawings showing the most characteristic phase of the position and movement of the speech organs in the production of the speech sound to be learned; (2) a carefully worded paragraph describing the most characteristic phase of the position and movement of the speech organs in the articulation of the speech sound was presented just below the drawings; (3) below the paragraph just mentioned came four key words most likely to be in a pupil's spoken vocabulary; and (4) twelve, one-syllable nonsense words were introduced just below the key words.

The instructional material was divided in three parts. Part I contained instructional material on the consonants. Part II contained instructional material on the vowels. Part III contained instructional material on blending syllables.

In Part I which contained instructional material on the consonants it was necessary to teach the vowels ē and ĭ. The vowels ē and ĭ were used as a nucleus in building one-syllable nonsense words for practice on blending the speech sounds represented by the consonants. The order in which the consonants were introduced was determined by a physiological approach to each consonant sound in which the mode of production was the basis of classification, and the order of presentation was as follows: p, b, t, d, k, g, m, n, ng, l, wh, w, f, v, th, ~~th~~, s, z, r, sh, zh, y, ch, j, and h. This order for the initial consonants in each nonsense word was adhered to without exception in the construction of the drills in the instructional material.

Part II contained instructional material on the vowels. The order of presentation of the vowels was determined by a physiological approach to each vowel sound, in which the position of the tongue was the basis of classification, namely: (1) high in the front of the mouth to low front; (2) high middle to low middle; and (3) high in the back of the

mouth to low back. Consequently, the order of the presentation of the vowels was as follows: ē, ĭ, ā, ĕ, â, ă, û or ẽ, ŭ, ȧ, o͞o, o͝o, ŏ, ô, o͡o, o͝o, and ȧ. The obscure or neutral vowel sound represented by ŭ, ȧ, ȧ, ŭ, ŭ, ŭ, and ŭ, was taught immediately following the ŏ symbol. It was necessary to make this exception because the obscure vowel sound occurs only in words of two or more syllables.

Length of Experiment.—This experiment was continued for fifty, consecutive school days. The pupils experienced only one drill period each day. Each drill period was twenty minutes in length.

The Data

Table I shows the mean grade placements made on The Metropolitan Achievement Tests which were administered after this experiment closed.

The data presented in Table I reveal that these pupils were making satisfactory progress in school as evidenced by the grade placements made on The Metropolitan Achievement Tests.

Table II shows the means made by Grades V, VI, VII, and VIII on The Word-Analysis Test, Form A. This test was administered at the beginning of this experiment and also at the close of the experiment.

Table II shows that the mean on the initial test was 1.02 words for the fifty pupils. Table II shows that the mean on the final test (one trial) was 18.38 words.

Table II shows that on the final test (one trial) Grade VIII analyzed correctly three words out of every five words attempted. Grade VII analyzed correctly four words out of every five words attempted. Grade VI analyzed correctly three words out of every five attempted. And, Grade V analyzed correctly four words out of every five words attempted.

The Summary and Recommendations

1. The pupils' mean grade placements made on The Metropolitan Achievement Tests seem to indicate that they were making satisfactory progress in school.
2. A mean of 1.02 words made on the initial test seems to indicate that the functional method is not very effective.
3. A mean of 18.38 words (on the first trial) seems to indicate that a drill program using carefully prepared instructional material and accompanied by the functional method is much more effective than the functional method used without a program of instructional material for drill on the pronunciation skills.

Table I

A Comparison of the Mean Grade Placements of Grades V, VI, VII, and VIII on the Metropolitan Achievement Tests

Grade	Mean Grade Placements				
	Reading	Language	Spelling	Arithmetic	Mean
8	8.5	9.2	7.9	9.1	8.9
7	8.4	8.8	8.3	8.5	8.6
6	7.2	7.4	6.8	7.6	7.1
5	6.8	7.0	6.5	6.3	6.5

Table II

A Comparison of the Means Made by Grades V, VI, VII, and VIII on the Word-Analysis Test, Form A

Grade	Mean Number of Words Analyzed Correctly			
	Initial Test	Final Test		
		First Trial	Second Trial	Third Trial
8	0.73	15.55	20.73	22.73
7	1.33	19.60	22.93	23.93
6	1.20	18.73	22.20	23.53
5	1.67	19.00	21.89	23.44
Mean	1.02	18.38	22.01	23.46

4. A mean of 18.38 words shows that these pupils analyzed correctly not quite four words out of every five words attempted.
5. The number of pupils used in this study make the data inconclusive; the data may be indicative of what one may expect when such instruction is placed in a school's instructional program.
6. A study should be conducted which would indicate the most effective complete symbolization to use as a phonetic aid in analyzing new words which pupils cannot pronounce from memory.
7. An experiment should be run which would reveal the best method to use in teaching the pronunciation skills.
8. Research which would show the best type of instructional material to use in teaching the pronunciation skills is needed.
9. Someone should direct a study which would indicate the most desirable grade placement at which to teach the pronunciation skills.

CONSTRUCTING EQUIVALENT SALARY SCHEDULES

JOSEPH B. HILL
764 Twentieth Ave., San Francisco, California

In the construction of a salary schedule for teachers there is wide range for choice as to:

1. The number of annual increments to be included in the schedule
2. The plan of increment distribution
3. The amount of minimum salary in relation to the maximum.

Thus salary increases may be spread over a greater or lesser portion of the teaching career; increments may be made equal, or they may be so arranged as to vary in amount in accordance with any desired plan; and the minimum salary may be set at a low figure with the resultant maximum correspondingly high, or it may be fixed at a relatively high figure with a lower maximum. By varying the choices made in these respects, salary schedules differing widely in pattern may be constructed.

It is a generally accepted principle that salary schedules should be built so as to attract and hold good teachers. There is, however, no agreement as to the salary schedule pattern or patterns that are, for a given expenditure, most effective in making that principle operative.

In the study of salary schedule pattern it is desirable to have at hand a method for constructing salary schedules that differ in pattern yet represent the same level of expenditure for teaching service. Such a method will be described here. The term "equivalent" will be used as descriptive of the required schedules.

Differences between salary schedule patterns are essentially differences in time distribution of payments. In arriving at a means for equalizing such differences two factors, interest and length of teacher service, must be taken into account. These factors will be considered separately.

I. THE INTEREST FACTOR

A payment made now on one salary schedule but on another deferred to a later time must, if it is to have the same value on both schedules, be increased in the second case by the amount which would accrue to it if it were placed at interest during the time of deferral. This is true from the standpoint of the teacher who is to receive the payment, since he does not in the meantime have the use of the amount involved; it is equally true from the standpoint of the school district which is to make the payment, since the amount deferred remains available for use during the intervening time by the district or by its individual taxpayers. Thus if, for example, two annual salary payments of one thousand dollars each are to be replaced by two others of which the first is eight hundred dollars, the second payment must include, in addition to the remaining twelve hundred dollars, the interest on two hundred dollars for one year.

The concept of present value is a useful one in reducing to comparable terms payments made at different times. The present value of a sum of money to be paid at a future date is the value of that sum discounted at a given rate of interest to the present time. Conversely it is the amount that, if placed at compound interest now, would accumulate to the required sum at the payment date. The more remote the date at which a payment is to be made the less is the present value of the payment.

The present value of a series of payments is the sum of the present values of the individual payments. The present value at entrance of the payments provided by a salary schedule for a lifetime of teaching service is thus a measure that takes into account not only the amounts of the various salary payments but the point in the professional career at which each is made. Accordingly, allowance for the interest factor in constructing salary schedules of different patterns may be made by so arranging the schedules that the payments on each of them, over a period of years representative of the teaching lifetime, have the same present value.

Under the plan to be followed here, amounts of increments under various

arrangements of increment distribution are calculated so that payments above the minimum salary under each arrangement shall have a present value equal to the present value of the same number of annual payments of one dollar each. The unit schedules thus developed represent a per-teacher expenditure level of one dollar per year above the minimum salary. They may, by multiplication, be used in deriving schedules with minimum salaries as desired and conforming to any stated expenditure level.

In developing the necessary procedure the following symbols from investment mathematics are used:

$a_{\overline{n}|}$ (read a angle n): the present value of a series of n payments of one dollar each, calculated at a specified rate of interest and with each payment made at the end of the corresponding payment interval.

v^n: the value of one dollar discounted at a given rate of interest for n interest periods.

Values of the quantities represented by $a_{\overline{n}|}$ and v^n at various rates of interest are published in books on investment mathematics.

In expressing the present value at the time of the first payment of a series of annual payments of one dollar each, it is to be noted that the first payment is subject no discount and hence has a present value of one dollar. The present value of the remaining payments is $a_{\overline{n}|}$, n being the number of payments in the series *after the first payment*. The present value of the entire series, then, is $a_{\overline{n}|} + 1$, where n is one less than the total number of payments.

The present value of the payments above the minimum on a salary schedule that provides for a number of equal annual increases until the maximum salary is reached, then annual payments of the amount of the maximum, is given by the expression

$$\frac{W}{i}\left[(1+i)\,a_{\overline{n}|} - h\,v^n\right].$$

Here W is the annual increment, h is the number of increments to be allowed, i is the interest rate, and n, as before, is one less than the number of annual payments to be included in the schedule.[1] If this expression be set equal to $a_{\overline{n}|} + 1$ and the desired values be assigned to h, n, and i, the resulting equation can be solved for W. With W, the amount of the annual increment, thus made available, the payments above the minimum salary for each year of the schedule can be calculated. The equation may be conveniently written thus:

$$i(a_{\overline{n}|} + 1) = W\left[(1+i)a_{\overline{h}|} - h\,v^n\right].$$

The period of years chosen as representative of a teaching lifetime is to be determined on the basis of entrance and retirement conditions in the school district under consideration. If, for example, requirements for certification are such that the usual age of entering teachers is twenty-five years, and if the age of compulsory retirement is sixty-five years, the normal teaching lifetime in the district is forty years.

Varying the Number of Increments

Suppose that a salary schedule covering a forty-year period of service and containing ten equal increments is to be constructed, and that money is worth three per cent. Then in the present-value equation

$$i(a_{\overline{n}|} + 1) = W\left[(1+i)a_{\overline{h}|} - h\,v^n\right]$$

$n = 39$, $h = 10$, $i = .03$.

From tables, $a_{\overline{10}|}$ at $3\% = 8.530203$, $a_{\overline{39}|}$ at $3\% = 22.808215$, and v^{39} at $3\% = .315754$. Substituting these values in the equation and solving for W, we obtain $W = .12690$. Thus the required salary schedule, in terms of a per-teacher expenditure of one dollar per year above the minimum salary and unadjusted for length of teacher service, will contain ten increments of \$.12690 each. The resulting maximum salary will be $10 \times$ \$.12690, or \$1.269, above the minimum.

If the other conditions are held the same but the number of increments set at twenty, then $h = 20$, $a_{\overline{h}|} = 14.877475$, and the annual increment is found to be \$.07928. The maximum salary thus is $20 \times$ \$.07928 or \$1.586 above the minimum. If thirty is taken as the number of increments, the annual increment is \$.06665 and the corresponding maximum is \$2.000 above the minimum. If equal increases are provided every year after the first throughout the entire service lifetime, the annual increment is \$.06390 and the payment for the fortieth year is \$2.492 above the minimum salary.

These results are summarized in Table I. Figure 1 gives a graphical comparison of the four schedules after they have been adjusted for length of teacher service.

[1] The mathematical derivations and the calculations basic to the material in this article appear in: J. B. Hill, *The Construction of Equivalent Salary Schedules for Teachers*. Unpublished Doctor's Dissertation, University of California, Berkeley. 1940.

Varying the Plan of Increment Distribution

In a salary schedule where progress from minimum to maximum salary is not uniform but is made in increments of unequal amounts, the series of unequal increments may be considered as two or more consecutive series of equal increments. A classification of salary schedules according to the number of series of equal increments contained in each follows. In it, identifying type numbers are assigned to different forms of increment distribution. The basic present-value equation applying to each group is given as the group is listed.

Group A. Schedules Containing One Series of Equal Increments

Type 1. Increments equal in amount.

Here the equation is

$$i(a_n + 1) = W[(1+i)a_h - h\,v^n]$$

The use of this equation has already been illustrated.

Group B. Schedules Containing Two Series of Equal Increments

Type 2. Increments small at first, then larger.

Type 3. Increments large at first, then smaller.

The equation for Group B is

$$i(a_n + 1) = W[(1+i)(k_g\,a_g + v^g\,a_h) - v^n(g\,k_g + h)].$$

Here g and h respectively represent the numbers of increments in the two series, and k_g is the ratio which increments of the first series bear to those of the second. Thus if a schedule of Type 2 is to contain eight equal increments followed by twelve equal larger ones, and if each increment in the series of eight is to be one-half the amount of those in the series of twelve, then $g = 8$, $h = 12$, $k_g = .5$. If the schedule is to be of Type 3 instead, and if the increments of the first series are to be, for example, double those of the second, then $g = 8$, $h = 12$, $k_g = 2$.

Here, as in subsequent equations, the basic increment W is the increment of the final series—that is, the series which ends at the maximum salary—and all other increments are expressed as multiples or fractions of it.

Group C. Schedules Containing Three Series of Increments

Type 4. Increments small at first, then larger, then still larger.

Type 5. Increments small at first, then larger, then smaller.

Type 6. Increments large at first, then smaller, then still smaller.

Type 7. Increments large at first, then smaller, then larger.

The equation for Group C is

$$i(a_n + 1) = W[(1+i)(k_f\,a_f + v^f\,k_g\,a_g + v^{f+g}\,a_h) - v^n(fk_f + gk_g + h)].$$

Here f, g, and h in that order represent the numbers of increments in the three successive series, and k_f and k_g represent the ratios of the increments of the first two series, respectively, to those of the third series.

This classification can be continued as far as desired, with each successive group containing twice as many types as the one preceding. Thus Group D, composed of schedules with four series of increments, includes eight types.

Only one of the types of Group D will be used to illustrate the group. It will be called Type 8. In it increments of the first series are small, those of the second series larger, and those of the third and fourth series each successively smaller.

The equation for any group after the first may be obtained from the one for the preceding group by following these steps:

1. Assign a symbol to represent the number of increments in the added series of the new group, and another to express the relationship to the basic increment W of the increments of the added series. Suppose these symbols to be x and k_x respectively.

2. Write the equation for the preceding group, making these changes in the right-hand member:

a. In the coefficient of $1 + i$ multiply each term by v^x, then add the new term $k_x\,a_x$.

b. To the coefficient of v^n add the term $x\,k_x$.

Thus for Group D, using e and k_e as the added symbols, the required equation is:

$$i(a_n + 1) = W[(1+i)(k_e a_e + v^e k_t a_t + v^{e+t} k_g a_g + v^{e+t+n} a_h) - v^e(ek_e + fk_t + gk_g + h)].$$

Illustrative schedules of Types 1 to 8 are summarized in Table I and shown graphically in Figure 2.

II. The Service Factor

If all teachers remained in service for the full teaching lifetime, salary schedules equal in present value of payments would also be equal in the expenditure level represented. However, of a group of entering teachers some may be expected to drop out the first year and still others each successive year, with only a fraction of the original group remaining to receive the higher salary payments scheduled for the late years of service. Each teacher who does drop out will in general be replaced by another at or near the minimum salary, with a consequent saving to the school district. Thus in constructing a salary schedule in terms of a pre-established expenditure level, allowance must be made for the savings to be expected on account of teacher withdrawals.

The adjustment required to take account of the service factor involves the use of a service table, and is accurate only to the degree that the table used is applicable to the teaching group for which the adjustment is made. A service table for teachers sets up a hypothetical group of teachers at an early entrance age and follows them from year to year, showing for each year the expected number of separations from service—withdrawals, deaths, and disability retirements—as derived on the basis of experience in a particular district, state, or other area.

As applied to a schedule of payments above the minimum salary constructed in accordance with the method that has been described, the adjustment for length of teacher service may be made in six steps.

1. Note the number of teachers shown in the appropriate service table as in service at the entrance age contemplated in the schedule. Then for each year of service calculate the ratio to that number of the number of teachers expected, according to the table, to remain in service. Write as decimals the ratios so calculated.

The ratio for the first year will be 1. Denote that for the nth year after the first by t_n. Then of a given group of teachers who enter service at the entrance age, the fraction t_n is expected to remain in service during the nth year after the first.

2. Service tables give separate figures for men teachers and women teachers. So that the adjusted schedule shall apply to a teaching group composed of both men and women, calculate the values of t_n for each sex separately, weight the values in proportion to the number of each shown by experience as likely to be employed, and take the weighted average for each year as the values of t_n to be used in subsequent steps.

3. Find for each year the product of t_n and v^n, and add the amounts so obtained. This sum may be written

$$\sum_{n=0}^{n=n} t_n v^n.$$

If a group of teachers entering now were to be paid at the rate of one dollar per year each above a stated minimum salary, this sum would be the per-teacher present value of the expected payments to them above the minimum. Allowance is made in it for withdrawals, deaths, and disability retirements at the rates given in the service table.

4. On the schedule to be adjusted let y_n represent the payment above the minimum salary for the nth year after the first. The present value at entrance of each such payment is then $y_n v^n$. Find the products $y_n t_n v^n$ and add them. The sum may be written

$$\sum_{n=0}^{n=n} y_n t_n v^n.$$

If a group of teachers entering now were to be paid according to the schedule under adjustment, this sum would be the per-teacher present value of the expected payments to them above the minimum salary. Allowance is made in it for separations from service in accordance with the rates in the service table.

In arranging the calculations for this step, it is evident that the required products $y_n t_n v^n$ may be obtained by multiplying by the respective values of y_n the values of $t_n v^n$ found in the preceding step.

TABLE I

PAYMENTS ABOVE THE MINIMUM ON ILLUSTRATIVE UNIT SALARY SCHEDULES

Each schedule represents an annual per teacher expenditure of one dollar per year above the minimum salary. The following conditions are assumed throughout:

Interest rate: three per cent
Teaching lifetime: forty years
Service rates: those for San Francisco teachers, 1933–1939
Ratio of men to women teachers: one to four

Salary Schedule Type	Number of Increments	Series of Increments	Increment Ratios	Schedule Unadjusted for Length of Service		Service Coefficient (S)	Schedule Adjusted for Length of Service	
				Increments	Maximum		Increments	Maximum
1	10	1	--	10 × $.12690	$1.269	1.062	10 × $.1347	$1.35
1	20	1	--	20 × .07928	1.586	1.096	20 × .0869	1.74
1	30	1	--	30 × .06665	2.000	1.132	30 × .0754	2.26
1	39	1	--	39 × .06390	2.492	1.141	39 × .0729	2.84
2	20	2	1:2	10 × .05765 10 × .11530	1.730	1.120	10 × .0646 10 × .1291	1.94
3	20	2	2:1	10 × .09760 10 × .04880	1.464	1.086	10 × .1060 10 × .0530	1.59
4	20	3	1:2:4	6 × .03830 7 × .07659 7 × .15318	1.838	1.134	6 × .0434 7 × .0868 7 × .1737	2.08
5	20	3	1:2:1	6 × .05868 7 × .11736 7 × .05868	1.584	1.105	6 × .0648 7 × .1296 7 × .0648	1.75
6	20	3	4:2:1	6 × .12300 7 × .06150 7 × .03075	1.384	1.074	6 × .1320 7 × .0660 7 × .0330	1.49
7	20	3	2:1:2	6 × .09617 7 × .04808 7 × .09617	1.587	1.098	6 × .1056 7 × .0528 7 × .1056	1.74
8	20	4	1:4:2:1	5 × .03880 5 × .15520 5 × .07760 5 × .03880	1.552	1.104	5 × .0428 5 × .1713 5 × .0856 5 × .0428	1.71

5. Divide the result of Step 2 by that of Step 3. The quotient may be called the service coefficient. Denote it by S. Thus

$$S = \frac{\sum_{n=0}^{n=n} t_n v^n}{\sum_{n=0}^{n=n} y_n t_n v^n}.$$

Here the numerator of the fraction represents, in terms of present value, the per-teacher expenditure above the minimum salary that the adjusted schedule is to require; the denominator represents the lower expenditure that would be required under the given schedule were no adjustment applied to it. Accordingly if each of the payments on the unadjusted schedule is multiplied by the quotient S, a schedule requiring the desired expenditure will be obtained.

6. Finally, then, multiply the basic unadjusted increment W by S, call the result W', and calculate payments on a new schedule using W' as the basic increment. The adjustment for length of teacher service will now have been completed. Under service conditions reflected in the table used, and assuming that each teacher who leaves is to be replaced by a beginning teacher at the minimum salary, the adjusted schedule will represent a salary expenditure for each year of teaching service of one dollar per year above the minimum salary.

Since the rate of separation from service becomes increasingly greater at successive age levels, it follows that the greater the degree to which a salary schedule defers payments toward the retirement age the greater will be the resulting adjustment and hence the higher the maximum salary made possible under a given expenditure.

The results obtained when various illustrative unit schedules are adjusted in accordance with the steps outlined are given in Table I, and the adjusted schedules are shown graphically in Figures 1 and 2. The service table used in the adjustment was that for San Francisco Teachers, 1933 to 1939. It was made available through the courtesy of Ralph R. Nelson, Secretary-Actuary of the San Francisco City and County Employees' Retirement System. In the calculations the ratio of men to women teachers was taken as one to four, and an entrance age of twenty-five years was assumed.

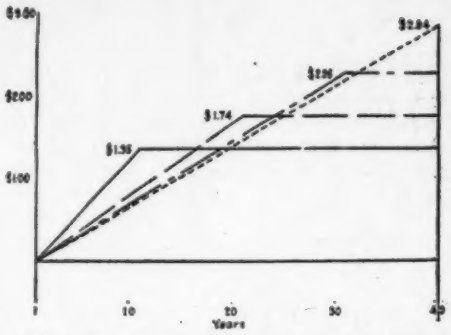

Figure 1.—Payments above the minimum on illustrative unit salary schedules differing in number of increments.

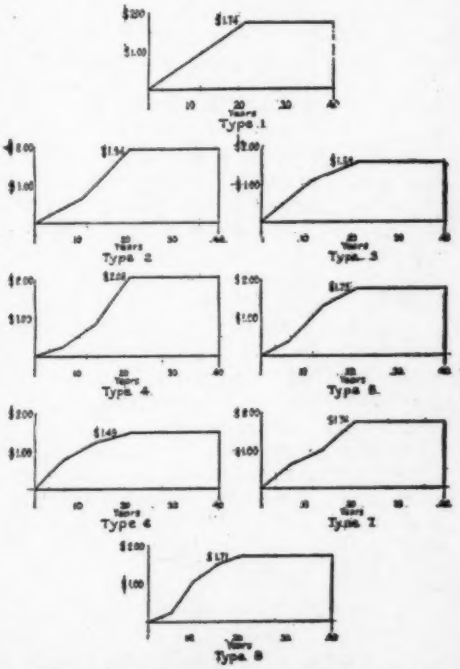

Figure 2.—Payments above the minimum on illustrative unit salary schedules differing in plan of increment distribution.

The withdrawal rates expressed in the San Francisco table are relatively low, expressing a high degree of stability in teacher personnel. In areas where teacher turnover is greater the service adjustment would be correspondingly increased.

Varying the Relation of Minimum Salary to Maximum

To construct equivalent salary schedules differing in relation of maximum salary to minimum, it is only necessary to vary the minimum salary while holding the expenditure level constant. To illustrate the procedure for this, let us use the twenty-increment Type 1 schedule of Table I and assume a desired expenditure level of $3000 per year for each teacher. Let us then determine the schedules resulting when $1500, $2000 and $2500 respectively are taken as the minimum salary.

From the table, the maximum salary on the unit schedule in question is $1.74 above the minimum for each dollar the expenditure level exceeds the minimum. For a minimum salary of $1500, the amount by which the expenditure level exceeds the minimum is $3000 — $1500, or $1500. The maximum salary, then, is $1500 + ($1500 × 1.74) or $4110.

With $2000 as the minimum, the expenditure level exceeds the minimum salary by $1000, and the maximum is $2000 + ($1000 × 1.74), or $3740. In the same way, a minimum salary of $2500 results in a maximum of $2500 + ($500 × 1.74), or $3370.

These three schedules are compared graphically in Figure 3. Any desired amount might be taken as the minimum salary, provided only that it be not greater than the expenditure level to be maintained.

Differences in salary schedule pattern as to number of increments, plan of increment distribution and relation of minimum salary to maximum have been illustrated separately for greater clarity. However, any combination of these kinds of variation is possible

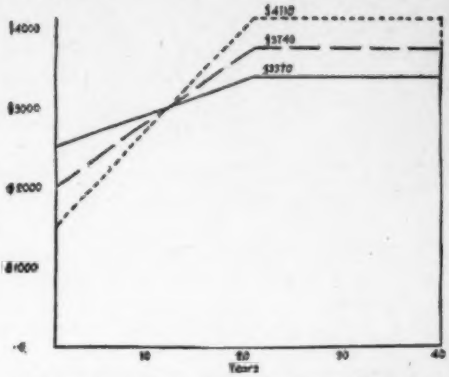

Figure 3.—Illustrative salary schedules differing in relation of minimum salary to maximum, constructed to represent a per-teacher salary expenditure of $3,000 annually.

under the method that has been described. For example, a schedule with high minimum salary and few increments decreasing in amount may be constructed equivalent to one with low minimum salary and a large number of increments increasing in amount, with consequent wide divergence in scheduled payments and in maximum salary.

No attempt has been made here to evaluate the various salary schedule patterns described, or to establish or recommend any salary level. The problem attacked has been solely one of method, and the schedules included have been for purposes of illustration only. A procedure has been described for the construction of salary schedules differing in pattern yet representative of the same level of salary expenditure. Further work is needed to determine the relative effectiveness of various schedule patterns in attracting and holding good teachers.

Euck